Building stock energy modelling

Brøgger, Morten

DOI (link to publication from Publisher):
10.5278/vbn.phd.eng.00077

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Brøgger, M. (2019). *Building stock energy modelling*. Aalborg Universitetsforlag. Ph.d.-serien for Det Ingeniør- og Naturvidenskabelige Fakultet, Aalborg Universitet https://doi.org/10.5278/vbn.phd.eng.00077

BUILDING STOCK ENERGY MODELLING

BY
MORTEN BRØGGER

DISSERTATION SUBMITTED 2019

AALBORG UNIVERSITY
DENMARK

Building stock energy modelling

Ph.D. Dissertation

Morten Brøgger

Aalborg University
Danish Building Research Institute
A. C. Meyers Vænge 15
DK-2450 Copenhagen

Dissertation submitted:	January 11, 2019
PhD supervisor:	Kim B. Wittchen Danish Building Research Institute Aalborg University in Copenhagen
Assistant PhD supervisor:	Ole Michael Jensen Danish Building Research Institute Aalborg University in Copenhagen
PhD committee:	Senior Researcher Jørgen Rose (chairman) Aalborg University Professor Paul Ruyssevelt University College London Professor Carsten Rode Technical University of Denmark
PhD Series:	Faculty of Engineering and Science, Aalborg University
Department:	Danish Building Research Institute

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-380-8

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2019

Preface

With increasing global warming, there is a need for political interventions that will reduce greenhouse gas (GHG) emissions. As resources are limited, it is crucial to prioritise measures that will lead to the largest emission reductions.

Reducing energy consumption in buildings has been identified as a key area in mitigating GHG emissions, with energy-efficiency improvements in buildings often being cost effective.

The present study considers building stock energy modelling as a tool for identifying and illustrating potential energy efficiency improvements, as well as the cost-effectiveness of these, in the built environment. This includes a description of different modelling methodologies and data used in building stock energy models, as well as the development of a model that seeks to advance specific aspects of building stock energy modelling.

The present dissertation was submitted for the degree of Doctor of Philosophy in January 2019 after three years of studies at the Danish Building Research Institute, Aalborg University in Copenhagen.

I hope you will enjoy reading my thesis.

Morten Brøgger
Aalborg University, January 11, 2019

Abstract

Reducing energy consumption in buildings is crucial in order to mitigate climate changes, as energy consumption in buildings constitute a large fraction of the world's total energy consumption. However, due to the diversity of buildings, as well as how these are used, assessing the energy efficiency of a building stock (be it a district, a city or an entire country), as well as how it may be improved cost-effectively, is a complicated matter. Moreover, access to energy relevant data is often scarce, e.g. for privacy reasons, complicating things further. However, with policies being put in place for monitoring the energy efficiency of buildings, increasing amounts of data are becoming available.

Building stock energy models provide the means for assessing and quantifying the energy-saving potential of a building stock, as well as for displaying it to different stakeholders. With access to different data, different models can be set up. However, common to all models is that they must be capable of serving different stakeholders' needs and that they must provide reliable results. In addition, it should be possible to assess the trade-off between investments in energy efficiency and other investments (e.g. investments in renewable energy sources or improvements of the energy supply system), e.g. through integration with other models.

In the present study, a sample-based hybrid bottom-up building stock energy model was developed for assessment of the energy-saving potential of the Danish residential building stock. Building-physical data from the Danish Energy Performance Certificate (EPC) database was used in combination with billed energy consumption data for setting up the model.

Following a review of building stock energy modelling methods, including an assessment of advantages and disadvantages, a sample-based (building-physical) model was set-up. By calculating the energy demands of each building individually it was possible to emulate the heterogeneity of the building stock nicely. Moreover, this made the model exceedingly versatile as any subset of the building stock could be studied without fitting the model specifically to each segment.

Next, the unique building-physical models were combined with energy consumption information in a statistical setting to form a hybrid model. This model made it possible to assess the accuracy of the model, as well as to calibrate it on a whole-building level. Moreover, the hybrid building stock energy model offered a nice interpretation that could be used for assessing rebound effects as well as how these differed among buildings with different characteristics.

Finally, the design of the model facilitated easy integration with other models, in this case an end-user economic model, in order to assess synergies and trade-offs between investments in energy efficiency and other investments in different contexts.

Resumé

Bygningers energiforbrug skal nedbringes, hvis de verserende klimaforandringer skal imødegås, da energiforbrug i bygninger tegner sig for en stor del af verdens samlede energiforbrug. Diversiteten af bygningsmassen (hvad enten der er tale om et område, en by eller et land), såvel som måden bygninger bliver brugt på, gør det imidlertid kompliceret, at kortlægge energieffektiviteten af bygningsmassen of videre hvordan denne kan forbedres. Ydermere er adgangen til energirelaterede data ofte begrænset, blandt andet af hensyn til privatlivets fred. Dog er flere data blevet tilgængelige i forbindelse med indførelsen af lovpligtig monitorering af bygningers energieffektivitet.

Bygningsmasse-energimodeller kan bruges til, at skabe et overblik over energieffektiviteten af en bygningsmasse, samt det dertilhørende energisparepotentiale. Hvilken type model der kan opstilles afhænger af, hvilke data, der er til rådighed. Et fælles krav til alle modeller er imidlertid, at de skal være alsidige, for at kunne imødekomme forskellige aktørers behov, samt at de skal levere pålidelige resultater. Herudover skal modellen gøre det muligt, at sammenholde investeringer i energiforbedringer med andre investeringer f.eks. investeringer i vedvarende energiteknologier eller energieffektivisering af energisystemet, eksempelvis gennem integration med andre modeller.

I det nærværende studie blev en stikprøve-baseret hybrid bygningsmasse-energimodel udviklet til fastlæggelse af besparelsespotentialet ved energirenovering af den danske boligmasse. Til dette formål blev bygningsfysiske data fra den danske energimærkningsdatabase anvendt i kombination med registrerede energiforbrug.

Efter en gennemgang af metoder til opsætning af bygningsmasseenergimodeller, herunder kortlægning af fordele og ulemper ved de enkelte modeltyper, blev en stikprøvebaserede (bygningsfysiske) model udviklet, til beregning af hver enkelt bygnings energibehov. Denne måde at modellere bygningsmassen på bød på to uovertrufne fordele; for de første var det muligt at emulere forskelligheden (dvs. heterogeniteten) af alle bygninger, for det andet gjorde modellen det muligt at studere en hvilken som helst delmængde

af bygningsmassen, uden at opstille separate modeller af hvert segment.

Denne unikke bygningsfysiske beskrivelse blev herefter kombineret med energiforbrugsdata i en statistisk (hybrid) model, hvilket gjorde det muligt, at vurdere nøjagtigheden af modellen, samt at kalibrere den på bygningsniveau. Ydermere gjorde fortolkningen af hybrid-modellen det muligt, at estimere rebound-effekter, samt hvordan disse varierer mellem bygninger med forskellige karakteristika.

Slutteligt gjorde denne metode til at modellere bygningsmassen på det muligt, at integrere modellen med andre modeller, i dette tilfælde en privatøkonomisk model, således at synergieffekter og kompromiser mellem investeringer i energioptimering og andre investeringer kunne vurderes i forskellige sammenhænge.

Acknowledgements

This thesis was sponsored by the Danish Innovation Fund (Innovationsfonden) [grant number 4106-00009A] and Aalborg University. All work should be seen as a contribution to the SAVE-E project [11], as well as the International Energy Agency (IEA) Annex 70 [38] and the IEA Annex 71 [25].

I would like to thank my supervisors, Kim B. Wittchen, Ole Michael Jensen and Toke Haunstrup Christensen for assisting me and giving me valuable advise.

Likewise, I would like to thank my colleagues at the Danish Building Research Institute for supporting me with anything and everything. A special thanks to 'the grumphy old men'; it has been a pleasure working with you.

I would also like to thank my co-authors, Peder Bacher and Henrik Madsen, as well as Henrik Klinge Jakobsen and Mattia Baldini; it was a pleasure working with you and get your perspective on our common challenges.

Finally, I would like to thank my family for supporting me by taking an interest in my work. A special thank to my love and life companion Christine; thank you for believing in me and always supporting me. There is nobody I would rather have by my side.

List of publications

The following papers constitute the basis of this dissertation. Parts of the papers are used directly or indirectly in the extended summary of the thesis.

[I] Brøgger, m., Wittchen, K.B., Estimating the energy-saving potential in national building stocks - A methodology review, *Renewable and Sustainable Energy Reviews*, Vol. 82(2018), pp. 1489–1496, 2018.

[II] Brøgger, M, Wittchen, K.B., Flexible building stock modelling with array-programming *Proceedings of the 15th IBPSA Conference*, 2017.

[III] Brøgger, M., Bacher, P., Wittchen, K.B., A hybrid modelling method for improving estimates of the energy-saving potential of a building stock, *Energy and Buildings*, submitted, 2019.

[IV] Brøgger, M., Bacher, P., Madsen, H., Wittchen, K.B., Estimating the influence of rebound effects on the energy-saving potential in building stocks, *Energy and Buildings*, Vol. 181(2018), pp. 62–74, 2018.

In addition to the four papers above, a fifth paper has been initialised. However, as this paper is still only a draft, the content is still subject to change. Therefore, section 8 has been prepared as an abstract of the paper that can be read independently of the paper. This abstract represents only the author's own contribution to the paper, as indicated by the co-author statement. The paper has been included to illustrate one possible application of the model, which would otherwise be missing.

[V] Baldini, M., Brøgger, M., Jakobsen, H.K., Wittchen, K.B., Cost-effectiveness of energy conservation measures in a Danish district heating system, Draft, 2018.

In addition to the main papers, the following secondary publications have also been made.

[A] Brøgger, M., Wittchen, K.B., Energy Performance Certificate Classifications Across Shifting Frameworks, Proceedia Engineering, Vol. 161(2016), pp. 845–849, 2016.

[B] Brøgger, M., Wittchen, K.B., Quantifying Uncertainties in an Archetype-Based Building Stock Energy Model by Use of Individual Building Models, *World Academy of Science, Engineering and Technology International Journal of Energy and Environmental Engineering*, Vol:12, No:9, 2018.

It should be noted, the *World Academy of Science, Engineering and Technology* (WASET) has recently been announced a predatory journal[1] in Danish media. Therefore, the assessment committee is encouraged to read this paper extra carefully, as proper peer-review may not be ensured.

As part of the assessment, co-author statements have been made available to the assessment committee and are also available at the Faculty. The thesis is not in its present form acceptable for open publication but only in limited and closed circulation as copyright may not be ensured.

[1] A predatory journal is one whose aim it is just to make money without ensuring proper quality of peer-review prior to publication.

Contents

Papers 71

Nomenclature

$\bar{y}$	Average energy consumption of buildings in sample [kWh]
β	Regression coefficient
ϵ	Regression model residuals (error term)
$\hat{y}_i$	Predicted energy consumption of building i [kWh]
j	Fold number
k	Number of predictors in regression model
n	Number of buildings in sample (i.e. number of observations)
y_i	Billed (actual) energy consumption of building i [kWh]
BBR	Danish Building and Dwellings Register
BSEM	Building stock energy model *or* building stock energy modelling
CI	Confidence interval
CVRMSE	Coefficient of Variation of the Root Mean Square Error
DHW	Domestic hot water
DRY	Design reference year
EBC	Energy in Buildings and Communities Programme
ECM	Energy Conservation Measure
EPBD	Energy Performance of Buildings Directive
EPC	Energy Performance Certificate
ESP	Energy-saving potential
EUI	Energy Use Intensity [kWh/m2]

FSS	Forward Subset Selection
GHG	Green house gas
IEA	International Energy Agency
IPCC	Intergovernmental Panel on Climate Change
MAPE	Mean Absolute Percentage Error
MLR	Multiple Linear Regression
NMBE	Normalised Mean Bias Error
PEF	Primary energy factor
Q	Energy consumption [kWh]
RES	Renewable energy source
RMSE	Root Mean Square Error

Introduction

"You cannot escape the responsibility of tomorrow by evading it today."
- Abraham Lincoln

Extreme weather and climate refugees are just some of the consequences of global warming [27, 46]. Thus, limiting global warming is crucial in order to reduce climate changes as well as related risks [28]. Human activity is commonly recognised as a main cause of these climate changes in the scientific community [36, 37]. According the the Intergovernmental Panel on Climate Change (IPCC): "*Human-induced warming reached approximately* 1 °C *(*± 0.2 °C *likely range) above pre-industrial levels in 2017, increasing at* 0.2 °C *(*± 0.1 °C*) per decade*[2]" [28].

In order to mitigate these climate changes, an international agreement to keeping increases in global warming below 2 °C was adopted at the climate conference *COP21* in Paris in 2015 (commonly known as the *Paris Agreement*) [18]. In the European Union (EU), the *low-carbon roadmap 2050* has been put in place with a goal of reducing greenhouse gas (GHG) emissions by 80 % compared with 1990-levels [17]. This includes the *2020 climate and energy package* as well as the *2030 climate and energy framework* as intermediate targets to reduce GHG emissions, increase energy from renewable energy sources (RES) and improve energy efficiency by 2020 and 2030 respectively [15, 16].

A cornerstone in the target to increase energy efficiency is *the Energy Efficiency Plan*. Buildings play a major role in this plan, as buildings account for approximately 40 % of the total energy consumption in the EU, equivalent to 36 % of all GHG emissions [19]. In order to improve the energy efficiency of buildings, the *Energy Performance of Buildings Directive* (EPBD) has been put in place, promoting minimum requirements of new buildings, as well as existing buildings undergoing renovation, so far as these are cost-effective [13]. Moreover, the *Energy Efficiency Directive* (EED) requires European member states to energy upgrade 3 % of the buildings that are owned or occupied by public bodies[3] each year [14].

Reaching these targets in a cost optimal way requires a comprehensive overview of the building stock, including how energy efficiency may be improved cost-effectively. This is necessary in order to inform policymakers, as well as other stakeholders (e.g. utility companies, end-users, etc.), with respect to where investments are best placed. However, due to the multitude of building types that exist in a building stock, building stock energy modelling is a complex matter. Furthermore, as different stakeholders have different focuses, e.g. in terms of different subsets of buildings they think are interested in, a building stock energy model should be capable of addressing different questions and considering different subsets of the building stock.

With the increased focus on energy consumption in buildings, large amounts of data are being collected, e.g. through energy performance certificate (EPC)

[2] Pre-industrial being defined as the period 1850-1900

[3] i.e. 3 % of the heated floor area of building with a useful floor area above $500m^2$

schemes. This provides a potential source of information that could be used for assessing strategies towards reducing energy consumption in the built environment. However, despite the increasing data availability much data remains unavailable, e.g. personal information, for privacy reasons.

A wide range of building stock energy models have already been developed with the purpose of mapping paths towards a sustainable future. However, challenges remain in terms of encapsulating the complexity of the building stock, as well as providing accurate estimates of the energy-saving potential of building stocks and subsets hereof, which can be used for informing policy makers and other stakeholders.

Objectives and research question

The purpose of this study was to develop a versatile building stock energy model that could be used for estimating- and illustrating the energy-saving potential of the existing residential building stock in order to inform policy makers and other stakeholders with respect to investments in energy efficiency upgrades.

With this objective in mind, the following research question was formulated:

How can a versatile building stock energy model, which encapsulates the complexity of the building stock, be set up for accurate estimation of the energy-saving potential of a building stock or subsets hereof?

In order to answer this question, a review of existing modelling methodologies was conducted with the purpose of describing different aspects of building stock energy modelling as well as identifying short-comings in present-day models.

Secondly, a model that could be used for evaluating different energy-efficiency upgrade scenarios, while accounting for the identified short-comings, was developed in three stages.

Finally, the cost-effectiveness of a suite of energy-conservation measures was evaluated, using a subset of the Danish building stock as a case study.

Delimitation

In the present thesis, four residential building types with similar usage profiles were considered. However, the applied methods could also be used in other building types with minor modifications.

Furthermore, only existing buildings were considered, though not considering demolition. This entails that dynamics of the building stock, in terms of rates of new-built as well as demolition rates, were not considered.

In terms of data, only existing data sources were used. Hence, no in-field data was collected. Likewise, only available data was used. Therefore, only building related data (i.e. no occupant- or household related data sources) were used.

Finally, only buildings for which data was available were considered. This entails that the results were not extrapolated to the entire building stock. However, as most of the contribution of this dissertation is methodological (rather than applied), extrapolation could be seen as a natural next step in the development of a building stock energy model that could cover the entire building stock.

Context

Building stock energy models can be developed for a wide variety of usages depending on the context they are developed within. In buildings, energy is used for different purposes, at different times, depending on the use of the building. Therefore, analyses of building's energy consumption depend on the buildings under consideration.

This chapter briefly introduces *a building stock* as a concept and outlines what characterises a building stock energy model.

Secondly, it briefly introduces the Danish building stock, including the energy efficiency of the building stock. This serves as a rational basis for the subset of buildings that was considered in this thesis.

1 Characteristics of a building stock

A building stock often comprises tens of thousands of building of different types, which are used differently. Therefore, a building stock is extremely complex. Moreover, different stakeholders (e.g. politicians, utility companies and private end-users) are often involved.

Since a building stock can comprise an arbitrary number of buildings, the following definition of a building stock was adopted, in order to distinguish a building stock from individual buildings.

A building stock is defined as a collection of buildings too large to monitor in detail

In this context, 'monitoring' includes everything from building and socio-economic (occupant related) characteristics to indoor- and outdoor environmental properties, which could have an effect on energy consumption in buildings. These properties could potentially be measured for one or a few buildings; however, at larger scales it becomes extremely resource demanding [39]. Therefore, various data sources must often be compiled in order to access the energy performance of a building stock [43].

1.1 Characteristics of a building stock energy model

A building stock energy model (BSEM) provides the means to evaluate energy-efficiency strategies, as well as communicating results to non-experts. These models can operate at different scales, ranging from neighbourhoods over cities to national building stocks. Different types of BSEMs can be developed for different applications, differing in terms of data requirements as well as modelling capabilities.

Due to the extend of a building stock, all data necessary to model it in detail is rarely available. Moreover, the resource intensity of data collecting renders this option impractical or even impossible. Therefore, building stock energy models often are constructed on the basis of very different input data. In many cases, existing data sources can often be used in combination with normative assumptions. However, care must be taken when using assumptions as substitutes for missing data, as faulty assumptions could result in bias in the model(s). Different types of building stock energy data are discussed in the subsequent chapter along with potential data sources and data used in this study.

Dependent on the data that is available, different method can be used for setting up a BSEM, each of which have different advantages and disadvantages. section 5 assesses some of the leading edge methods used in building stock energy modelling as well as related pros and cons. Likewise, different methods exist for representing the building stock, which differ in terms of complexity, as well as flexibility.

The versatility of a BSEM is central to meet different stakeholders' needs, e.g. in terms of modelling different subsets of the building stock. Likewise, it is crucial to account correctly for the diversity of the building stock in order to ensure a correct representation.

Data scarcity may render it necessary to combine modelling techniques in order to correctly account for different aspects that could affect energy consumption. This includes validation of the model results, in order to ensure the credibility of the model.

Finally, it may be necessary to consider competing alternatives, when evaluating cost-effectiveness the energy-saving potential of a building stock. Therefore, integrability with other models is a desirable feature of a BSEM.

Despite differences in terms of input data and modelling techniques, all building stock energy models share some common characteristics. These include being complex, both in terms of the buildings under consideration and in terms of the users of the models (i.e. stakeholders). Furthermore, data access is often limited, making it even more difficult to estimate the heat consumption.

In summary, building stock energy modelling is characterised by:

- multiple actors and stakeholders
- limited access to detailed data
- a large degree of complexity

Therefore, building stock energy models must be able to encapsulate the complexity in order to communicate findings to non-experts. However, simplifications made in the model should not compromise the accuracy of the results to a degree that could cause wrong decisions to be made.

2 The Danish context

The Danish building stock, like many others, is characterised by comprising many different types of buildings[4], which serve a multitude of purposes, which makes it complex. Also like in many other countries, the energy consumption in the Danish building stock accounts for a significant part of the

[4] At the time of writing, 28 different superordinate building categories were registered in the national Building and Dwelling Stock Register (BBR)

final total energy consumption. In 2016, the total energy consumption in Danish buildings amounted to 405 TJ, of which 196 TJ was used in households, 126 TJ were used in production facilities (agriculture, fishing, etc.) and 83 TJ were used in trade- and services (i.e. wholesale, retail, private- and public services) [6], see Figure 1

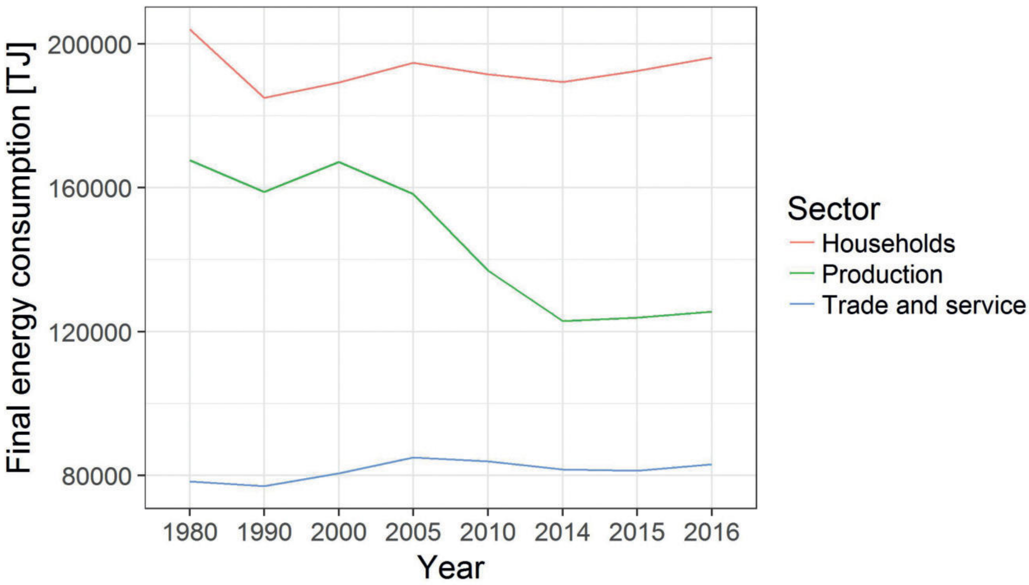


Fig. 1: Final energy consumption, i.e. delivered energy, in different sectors of the Danish building stock (climate corrected). All numbers were taken from the Danish Energy Statistics 2016 [6].

Despite (successful) efforts to improve the energy-efficiency of the building stock, it should be noted that the final (total) energy consumption has hardly decreased over the past 25 years, even though the energy efficiency has improved.

Moreover, due to the differences in use among the different sectors, analysing the total building stock would be beyond the scope of this thesis. Therefore, households (i.e. residential buildings) were chosen for analysis, as these account for the majority of the total energy consumption. In the present study, we considered the energy-efficiency, as well as the related energy-saving potential, of the four most common residential building types in Denmark[5]. These are farmhouses, detached single-family houses, terraced houses and blocks of flats. Examples of typical buildings of each of the four building types are shown in Figure 2.

[5]Dormitories (colleges), 24-hour care centres and other buildings for permanent residence, as well as secondary homes, were not included in the analyses

(a) Farmhouse (b) Detached single-family house

(c) Terraced house (d) Block of flats

Fig. 2: Examples of the four residential building types considered in this thesis

The distribution of the four building types is depicted in Figure 3

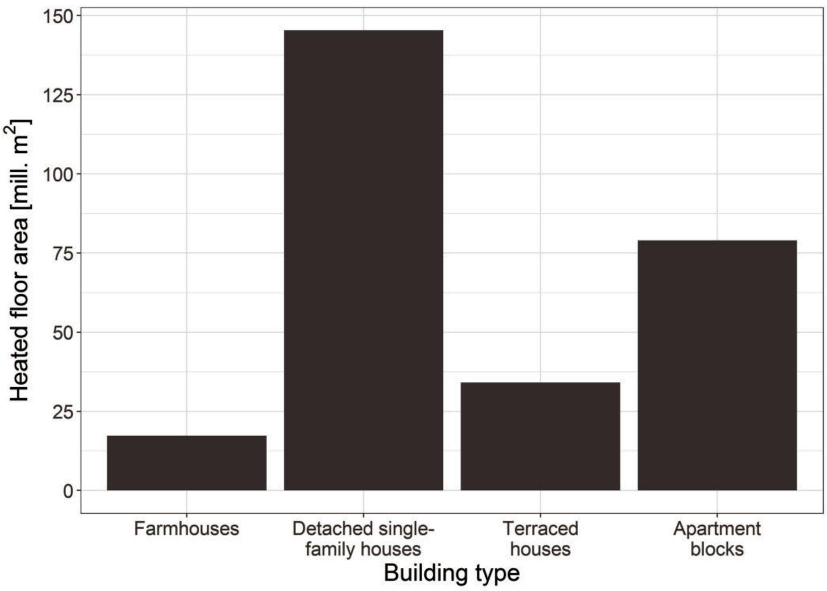


Fig. 3: Distribution of the four building types (adapted from **Paper B**)

Moreover, as heating (i.e. space heating and domestic hot water preparation) accounts for approximately 85 % of the heat consumption in an average Danish household [6], only heat consumption was considered in the present thesis.

Finally, only existing buildings were considered in the present thesis, due to the high energy efficiency and low rate of new built in the Danish building stock.

2.1 Energy efficiency of the Danish residential building stock

The long lifetime of buildings (some beyond 100 years) entails that a large part of the Danish building stock is very energy inefficient. Approximately 75 % of the Danish residential building stock was erected before the first energy-oriented tightening of the energy-related requirements in the Danish building code in 1979[6] [2], following the oil crisis in 1973-1974, see Figure 4[7].

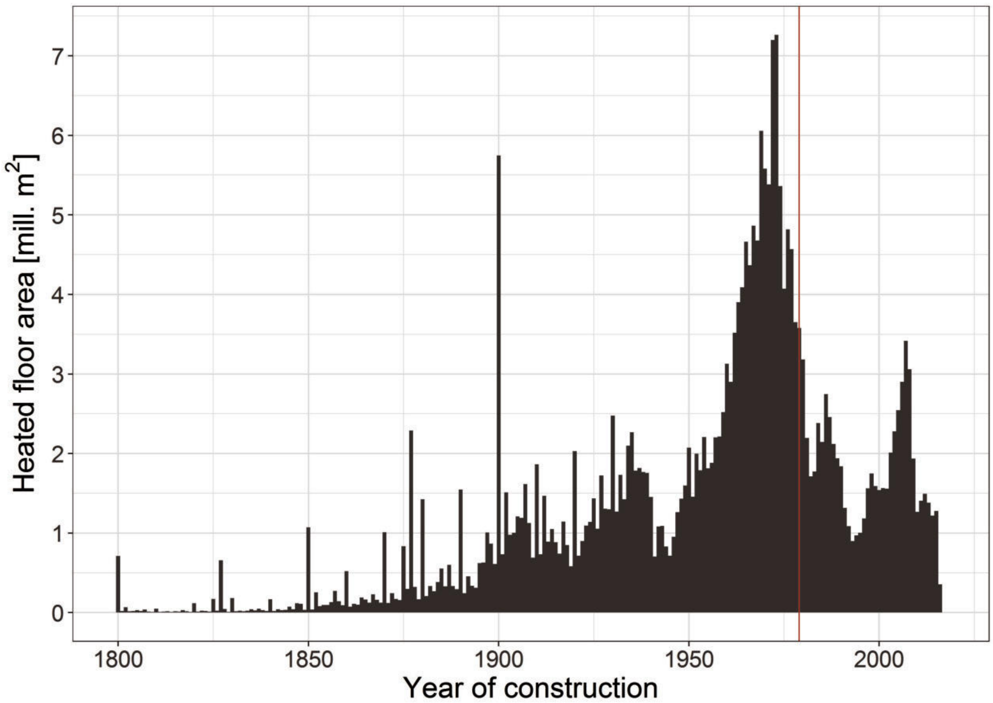


Fig. 4: The Danish residential building stock divided by year of construction. The red line denotes the year 1979 in which the insulation requirements in the Danish building code were tightened with a view to improve energy-efficiency.

However, due to the long lifetime, many buildings have undergone refurbishment, e.g. replacement of windows that have reached the end of their service life, thereby improving the energy-efficiency. Despite this fact, a large fraction of the residential building stock remains energy inefficient, see Figure 5.

[6]The first energy-related requirements in the Danish building code were imposed in the first national building code in 1961 with the purpose of preventing mould growth in building envelope elements.

[7]The large number of buildings erected in 1900 is likely a consequence of missing information in which case "old buildings" were assumed to be erected in 1900.

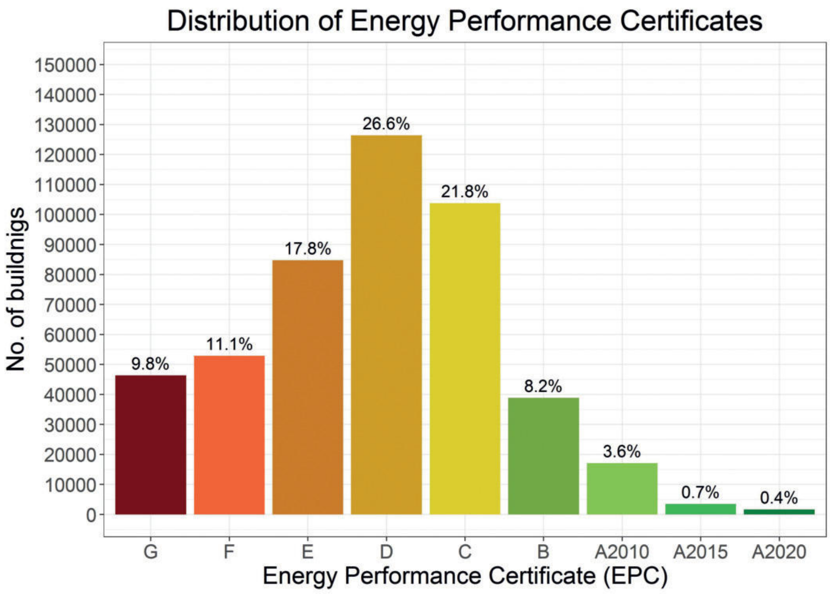


Fig. 5: Distribution of Energy Performance Certificates (EPCs) in the Danish dwelling stock. In October 2015, 475 419 residential buildings had a valid EPC.

In Denmark, Energy Performance Certificates (EPCs) are issued on the basis of an asset rating (i.e. a calculated energy demand). This includes heat demands for space heating and domestic hot water (DHW) preparation as well as electricity demands for building services[8]. In addition, primary energy factors (PEFs) are used for weighting different energy sources, in order to make them comparable[9]. The EPC evaluation scale for residential buildings is included in Table 1

EPC	**Threshold** $[\frac{\mathbf{kWh}}{\mathbf{m^2 \cdot year}}]$
A2020	≤ 27.0
A2015	$\leq 30.0 + 1000/A_{\text{floor}}$
A2010	$\leq 52.5 + 1560/A_{\text{floor}}$
B	$\leq 70.0 + 2200/A_{\text{floor}}$
C	$\leq 110 + 3200/A_{\text{floor}}$
D	$\leq 150 + 4200/A_{\text{floor}}$
E	$\leq 190 + 5200/A_{\text{floor}}$
F	$\leq 240 + 6500/A_{\text{floor}}$
G	$> 240 + 6500/A_{\text{floor}}$

Table 1: Danish EPC scale for residential buildings [5]

[8]Appliances and lighting is not considered in the EPC of residential buildings

[9]In the current building code (BR18), electricity has a PEF of 1.9, district heating has a PEF of 0.85 and other heating has a PEF of 1.0 [45]

The large share of energy-inefficient buildings suggests a significant energy-saving potential in the Danish residential building stock. Realising this potential requires a detailed overview of the energy-efficiency of the building stock, as well as the costs- and effects of implementing energy-conservation measures. However, this is complicated by the complexity of the building stock. Therefore, transparent and flexible methods are needed for facilitating relevant analyses that can help inform policy markers and other stakeholders with respect to investment decisions and political interventions (e.g. subsidy schemes).

Building stock data

"Errors using inadequate data are much less than those using no data at all."
- Charles Babbage

Building stock energy modelling is inextricably linked with the data that is available. Therefore, building stock energy modelling and building stock data should be seen in combination with one another [35, 40].

Access to good data has thus been recognised as a key prerequisite in building stock energy modelling [35]. However, gathering enough data to model the energy demand of a building stock is a resource demanding process [39]. Therefore, existing data sources, e.g. public registries, are often used for modelling the building stock. This includes compilation of several data sources [35].

However, as most public data sources have not been collected for the purpose of building stock energy modelling, relevant data is often missing. Moreover, data is often subject to privacy, limiting the accessibility of these data. In order to overcome these obstacles, much existing data can be used as surrogates for needed data in some cases whereas in others, specific modelling techniques must be used.

In the following, we cover different data types that are important in building stock energy modelling, as well as the link between these data and the different types of building stock energy models. Likewise, potential sources for obtaining this information are discussed.

Lastly, the different data types used in this study are discussed along with the data preparation process.

3 Building, occupant, energy and environmental data

Four data types are central in building stock energy modelling (BSEM); these include building data (i.e. the physical properties of the building including building services and indoor environmental conditions), occupant data (i.e. building operation) and the external environment/climate, as well as the energy consumption [26, 35], see Figure 6.

Fig. 6: Four data types used in building stock energy modelling: Building-, occupant-, energy- and environment data (courtesy: IEA EBC Annex 70 data registry [26])

3.1 Building data

Building data can be divided into building-physical characteristics and indoor environmental conditions. Physical characteristics include information about the building envelope and the technical installations in the building (i.e. building services and appliances and lighting). This includes control strategies of the various technical systems in the building, e.g. solar shading and ventilation systems.

Indoor environmental conditions comprise factors that affects a building's energy balance and that relate to occupant behaviour directly (e.g. indoor temperatures though set-point temperatures) or indirectly (e.g. internal heat loads).

Estimating the energy-saving potential due to an energy-upgrade requires a building-physical description (or an equivalent building-physical interpretation) of the building stock, regardless of the employed method. Therefore, building data is essential to building stock energy modelling. Fortunately, building data is becoming increasingly available through surveys [34] and public databases [8]. In cases where information is missing, this can be estimated from the age of the building in combination with historical building traditions and building codes [31].

Building-physical characteristics

In many countries, building data is collected in central databases through Energy Performance Certifications, providing a unique insight into the physical condition of the existing building stock. In the EU, the Energy Performance of Buildings Directive (EPBD) requires all member states to issue energy performance certificates (EPCs) for all buildings that are constructed (new buildings), sold or rented out (existing buildings) [13].

EPCs that are based on an asset rating (i.e. calculated heat demands) provide unique information about the physical characteristics of the building stock that can be used for formulating energy-efficiency strategies if collected and stored appropriately. In most of the EU-28 countries, asset rating (or asset rating in combination with operational rating) is used for issuing EPCs [4], providing a potential source of information.

In Denmark, EPCs are based on an asset rating. All information in automatically collected in a central database, providing access to the physical properties of all component that are used for calculating the EPC, see Figure 7.

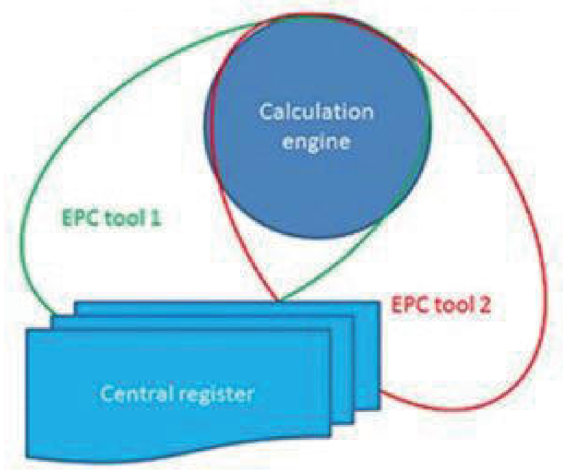


Fig. 7: Conceptual drawing of the collection and storage of data in the Danish EPC database (**Paper A**).

This way of collecting and storing data provides several advantages, including the possibility to exclude faulty data and re-calculating energy demands as discussed in **Paper A**.

Indoor environmental conditions

Indoor environmental conditions comprise several factors (e.g. indoor temperatures and air change rates) which interact with each other in complex thermodynamic relationships.

Common to these factors is that precise determination requires intrusive measurement which makes determination troublesome across large groups of buildings. Therefore, access to detailed indoor environmental conditions is often limited in building stock energy modelling.

In building-physics based BSEMs, normative assumption are often used in place of missing indoor environmental data. However, using the same assumptions across all buildings could turn out to yield erroneous (biased) results.

3.2 Occupant data

While energy consumption in buildings and occupant characteristics are not explicitly linked, they are linked implicitly (i.e. people do not demand energy, but rather a satisfactory indoor environment, which requires energy to uphold). As groups of people with similar characteristics tend to have similar indoor environmental preferences, occupant data often provide a useful proxy for the indoor environmental conditions, e.g. families with you children likely use energy differently than elderly people.

Typical occupant characteristics that are used in building energy modelling include age, income, education level, etc. However, due to data privacy constraints, occupant data are often difficult acquire at the individual building level. Moreover, usage profiles vary from day to day, making it difficult to determine average schedules. In addition, the household composition may change over time.

3.3 Energy consumption data

Energy consumption data plays a central role in BSEM. In statistical models, the energy consumption is used for fitting the model whereas in building physics-based models, the energy consumption is used for model validation.

Many utility companies collect registrations of metered/billed energy consumption that could be used in building stock energy modelling. However, the level at which this information can be acquired may also be limited by privacy constraints.

In Denmark, utility companies have been required to report registered consumption of district heating, natural gas and fuel oil back to the Danish Building and Dwelling Register (BBR) annually since 2011[10].

3.4 Environmental data

Environmental data include climate data that has an effect on the energy consumption in buildings, including (dry-bulb) air temperature, solar radiation

[10] Utility companies supplying less than ten end-users are excepted.

(direct and diffuse), humidity and wind-speeds from different directions.

Environmental (climate/weather) data are freely available for many locations, e.g. through the US Department of Energy (DOE) EnergyPlus homepage [10]. However, this information is often only available for larger area, i.e. on an aggregate level (e.g. by country or state), rendering it difficult to study effects of micro-climate.

4 Data used in this thesis

Two independent databases were used in the present thesis; these were the Danish Energy Performance Certificate (EPC) database and the Danish Building and Dwelling Register (BBR). The BBR is a publicly available, national register that contains information about every building in the country[11].

Information in the BBR includes individual building ID's, use, year of construction, size, (geo)location, energy supply and energy consumption, among other things. The responsibility to ensure that details are correct lies with the individual building owner. However, as the information is used as a basis of taxation, management of construction project as well as by other stakeholders, the correctness of of data is regularly being checked.

The information in EPC database, on the other hand, is reported by educated energy experts. Danish EPCs are building specific and require visual inspection by the energy expert[12]. In the EPC database 475,419 buildings were registered with a valid EPC. However, due to a large number of faulty data, the actual number of buildings considered decreased dramatically; see subsection 4.1. Moreover, because not all buildings in the BBR had a heat consumption associated with it, fewer buildings were available for some analyses. Lastly, it should be noted that data pre-processing (i.e. preparation) was an ongoing effort, for which reason the data that was used differed from one paper to another.

Because the EPC database and the BBR databases are independent of one another, combining the information could give an indication of which data were reliable and which were not.

Building data

Building-physical characteristics were retrieved from the EPC database. These included thermal properties (e.g. U-values) and specific sizes of each individual component in each building. This also included orientation of the building, as well as shadows on windows from surrounding obstacles. Figure 8 illustrates an example of how this data was stored.

[11] At the time of writing, 1,552,433 buildings, corresponding to 2,725,652 self-contained units, were registered in the BBR.

[12] in special cases, EPCs can be issued without a visual inspection, as well as for more buildings collectively.

	Identid	BuildingId	Component	Area	UValue	TempFactor
1	5114290	736669	Ceiling	23	0,6	1
2	5114291	736669	Ext.wall	10	2,1	1
3	5114292	736669	Ceiling	44	1,2	1
4	5114293	736669	Floor	50	0,3	1
5	5114294	736671	Ext.wall	200	1,8	1
6	5114295	736671	Ceiling	100	0,6	1
7	5114296	736671	Floor	100	1,2	1
8	5114297	736672	Floor	88	0,52	1
9	5114298	736672	Floor	10	0,34	0,7
10	5114299	736672	Floor	5	0,34	1,2

Fig. 8: Example of building-physical characteristics stored in the Danish EPC database (**Paper II**)

However, no data on the indoor environmental conditions in the considered buildings was available. Therefore, it was necessary to develop a model that could be used for estimation of the energy-saving potential while obviating the need for indoor environmental data.

Occupant data

No access was granted to any information about the occupants of the buildings in the present thesis. Therefore, it was necessary to develop a model that could account for user behaviour without modelling it explicitly; see section 7.

Energy consumption data

Energy consumption information was retrieved from the BBR database. This information was reported by utility companies in terms of metered (billed) consumption. Using metered data posses a potential thread, as all self-contained units (e.g. individual apartments) do not necessarily have their own meter. Moreover, registration periods do not necessarily span an entire year. Therefore, meter readings are annualised by an external company by means of the heating degree days in the registration period, to match the specific year (hence, the registered energy consumption is not standardised to a typical year). However, this entails that information about the registration period is lost; i.e. neither the season nor the length of the registration period was known.

Environmental data

In addition to the two aforementioned databases, environmental data in terms of the Danish Design Reference Year (DRY) was used[13]. The Danish DRY represents a *representative meteorological year* and is composed of monthly data from different years between 2001 and 2010, that are typical in terms of climatological variation [30]. These data were compiled from one of five-six (parameter specific) climatological zones based on the representativity and regularity of the observations [30]. Therefore, the environmental data set should be representative for the Danish building stock. However, local variations could occur, which could not be accounted for in this thesis.

Due to the way data was collected and stored, each individual building could be identified, i.e. information was available at the individual building level. This included information about the building-physical characteristics of the buildings as well as energy consumption registered by utility companies upon account.

Therefore, exploiting the full potential of the methods proposes here, it is necessary to have access to building physical, as well as energy consumption information, at the individual building level. However, this information could come in many forms and doesn't necessarily have to come from the same sources that data did in this study. Likewise, the proposed methods could be used with access to smaller amounts of data.

4.1 Data quality

The quality of input data is crucial in any model since the accuracy of the model output is limited by the accuracy of the model input. Therefore, data preparation constituted an important first step in the modelling process.

Despite the fact that EPCs were issued by trained energy experts, they were subject to two types of errors; these were:

- Misjudged values
- Faulty registrations (e.g. typing errors)

In general, misjudged values pose the biggest threat to the validity of the models, as these are prone to being systematic, e.g. systematic underestimation of U-values of external walls from a particular time period.

Both types of errors are subject to random errors, which pose a minor problem, as these tend to average out at scale. However, large one-directional errors also pose a problem, because the distributions of many input values are

[13] DRY is used for documenting that the of a building design complies with the legal requirements in the Danish building code

asymmetrical (i.e. skew). In building stock energy modelling, large positive values that are wrong could disturb the picture as many values cannot take on negative values, for simple physical reasons.
In order to eliminate large errors, a number of restrictions were imposed when extracting data from the two databases.

Data exclusion

The Danish EPC scheme has been criticised strongly in the media, because random checks have shown errors in 20 - 30 % of all EPCs [6, 7]. Therefore, limits were imposed on the basis of physical restrictions when extracting data, in order to discard faulty registrations. This sections describes the data exclusion process, including a description of potential errors that could not be detected.

As of 2015 (8 October 2015), the Danish EPC database contained data on 496.636 buildings[14]. With more than 30 related tables containing information about the building envelope and building services of each of these buildings, manual assessment of all data was not an option. Instead, a script was developed for removing apparent faulty registrations upon querying. This offered the advantage that the extensive amounts of data could be examined automatically; however, using an automated approach entailed that sub-categories could not be investigated in detail, i.e. only extreme values were discarded.

Criteria for exclusion

In **Paper II**, five criteria were outlined in order to uniquely identify a building; these were referred to as *essential information*. According to these criteria, a building must include an envelope, at least one window, be ventilated, have a DHW demand and include an internal heat load. It should be noted that these criteria were designed to match the structure of the Danish EPC database, e.g. windows with the same thermal properties with the same orientation may be merged into one window summing the area.

In addition to essential information, additional criteria were used for discarding building components that exhibited abnormal behaviour. These criteria were based on an extreme value approach, where registrations were discarded if they exceeded an upper- or lower limit defined for each part of the building. An example is given in Table 2.

[14] counting all registrations (rows) in the 'Building' table with *UseCode* 110 (farmhouses), 120 (detached single-family houses), 130 (terraced houses) or 140 (blocks of flats).

Characteristic	Criterion	No. of components	No. of buildings
Area [m^2]	> 0	18 786	6971
U-value [$\frac{W}{m^2 \cdot K}$]	$]0.03; 7]$	1545	1286
Temperature factor [-]	$[0; 1]$	12 160	11 273
Total		32 124	19 125

Table 2: Criteria for- and number of faulty registrations in the table containing building envelope elements in the Danish EPC database.

Indeed many faulty registrations appeared in the same building, or even in the same component, for which reason the total number of buildings discarded was smaller than the sum of flagged registrations. A full list of all criteria that were imposed upon extraction of data from the EPC database is enclosed in the appendix.

Upon removal of registrations that were deemed faulty, the building-physical description of 171 387 buildings were retained for analyses. It should be noted that the data quality assessment was an ongoing process, for which reason the same subset of buildings was not used in all papers.

In addition to the criteria imposed above, several other criteria could be used for excluding dubious registration, e.g. a maximum surface area to volume ratio. However, as such criteria would be building- and component specific, and since limits are ambiguous, this was beyond the automatic filtering approach used in this thesis.

In order to remedy the missing upper limits in the input values, additional filters were imposed after calculating building specific energy demands as addressed in each paper.

Building stock energy modelling

"All models are wrong, but some are useful."
- George E. P. Box

5 Modelling methods and aspects

*Unless otherwise specified, this section is based on **Paper I***

A building stock energy model (BSEM) can be used for creating an overview of the energy efficiency in a building stock as well as effects of imposing energy-conservation measures. BSEMs can be divided into two groups, namely top-down and bottom-up [29, 43]. These can be further sub-divided, as illustrated in Figure 9 (**Paper I**).

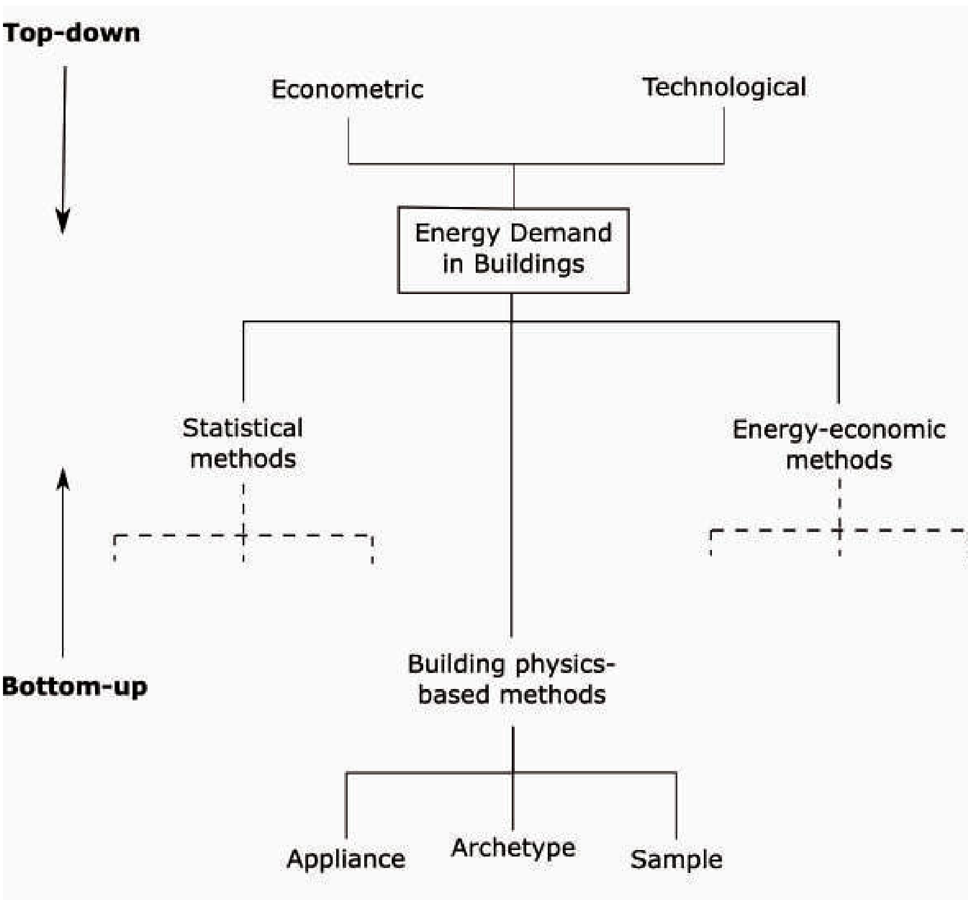


Fig. 9: Methods for estimating the energy end-use in buildings (**Paper I**)

Top-down models, which are characterised by relying on aggregate data, are ill-suited for estimating potential energy-savings [43]. Therefore, these were not considered in this thesis. In contrast, bottom-up models are characterised by relying on disaggregated data (e.g. physical characteristics of individual building components), making them suitable for estimating potential energy-savings [43].

Some of the general characteristics of different model types were discussed in **Paper I**, including advantages and disadvantages of different modelling approaches. This section summarises some of the pros and cons of different bottom-up models with a view to how current bottom-up models can be advanced.

5.1 Bottom-up building stock energy modelling

In general, bottom-up BSEMs can be divided into *statistical models* and *building-physical models*. In addition, *energy-economic models* exist, see Figure 9. These differ in terms of data requirements and modelling abilities, leaving each model type with different advantages and disadvantages.

Bottom-up statistical models

Statistical models, which are also known as data-driven models [1] or inverse models [32], rely on historical energy consumption data that is attributed to different end-uses or explanatory variables. Two features are characteristic to statistical models, namely the predictive performance and the interpretability of the model. However, these features are not necessarily compatible; i.e. the model with the best predictive performance may not offer the desired interpretation or vice versa. This is important to recognise, when setting up a statistical model. In addition, most statistical models rely on a number of underlying assumptions, e.g. randomness and independence of samples. Setting up a valid (and meaningful) statistical model requires that these assumptions be met. A distinct advantage of the statistical methods is that energy consumption can be estimated without having to specify user behaviour. However, purely statistics-based methods usually attribute energy consumption to a limited set of variables, making them inflexible in terms of evaluating a broad suit of energy-conservation measures [29].

Bottom-up building-physics based models

Building-physics based models (sometimes referred to as *engineering methods* [42], *first-principle* or *forward models* [1]) rely on building-, occupant- and environment data for calculating the energy consumption of each building. Building physics based models have been widely used for estimating the potential for energy-efficiency upgrades due to their explicit building component formulation. An inherent drawback of the building physical models is the need for specifying user behaviour, as data on indoor temperatures, air change rates, etc. are rarely available at the building stock level.

Building-physics based BSEMs can be subdivided into three categories according to (Swan et. al) [42], as illustrated in Figure 9. These are *Appliance-distribution models*, *Archetype models* and *Sample-based models*. Whereas appliance-distribution models seek to estimate the building stock energy consumption from appliance ownership and appliance ratings, archetype- and sample-based models estimate the whole-building energy consumption as a sum of individual contributions, i.e. energy balances.

Archetype-based BSEMs are presumably the single most used method for modelling the energy-saving potential in building stocks due to their simplicity and fairly limited requirements for data collection [39]. Building archetypes are composite (i.e. synthetic) buildings which are constructed on the basis of data that is representative of the building stock. Building archetypes are constructed by dividing the building stock into an arbitrary number of segments that have similar characteristics. A common segmentation of the building stock is by means of building type, construction period and climatic conditions, as suggested in the European projects TABULA and EPISCOPE [12, 44].

Despite its simplicity, the archetype approach is inflexible as a political decision-making tool, because archetypes must be defined prior to modelling, see section 6. Moreover, the number of archetypes scales proportionally with the number of characteristics one wants to include in the model. Table 3 shows the number of archetypes in relation to the information that is included in the model. Clearly, the number of archetypes increases dramatically with the information that is included in the model, thereby reducing the simplicity of the model.

Information	Building type	Construction period	Primary supply	Secondary supply		All
No. of segments	4	36	180	900	...	No. of buildings

Table 3: Number of segments depending on the information included in an archetype model using the Danish residential building stock as an example

Therefore, many traditional building archetypes only include few building characteristics, in order to keep the model simple. However, this could come at the expense of the variability in the model results, see subsection 6.2.

As opposed to the archetype approach, sample-based BSEMs calculate the energy demand of each building individually, thereby providing a more detailed picture of the energy consumption in the building stock. Thus, all information is included in the model, thereby representing the finish possible granularity, see Table 3. However, due to the level of detail, sample-based models require access to large amounts of data [42], which also makes these model computationally more intensive than archetype models.

In addition to statistical- and building-physics based BSEMs, energy-economic models exist. These make use of historical data on equipment use and efficiency and could thus be categorised as *Appliance* models under the *building physics* branch.

Finally, *hybrid models*, which draw on aspects from two or more model types, have been proposed in order to overcome the drawback of the traditional methods. One example of a hybrid model is the Canadian hybrid residential end-use energy and emissions model (CHREM), which used the output of a neural networks (statistical) model as input in a building physics based model, in order to model user behaviour (in terms of domestic hot water and appliances and lighting) correctly [41, 42].

5.2 The Performance Gap in building stock energy modelling

Due to the large number of factors that influence the energy consumption in buildings, many uncertainties persist. Therefore, several studies have encountered a discrepancy between the expected (i.e. modelled) energy demand of buildings and the actual energy consumption. This is often referred to as *the Performance Gap*. In particular, energy consumption appears to be linked with the building-physical energy performance of the building, in terms of *the Rebound Effect*[15]. Failing to recognise these effects were found to impact the estimated energy-saving potential immensely, in particular causing an overestimation of the potential.

The existence of the Performance Gap raises two questions, namely what causes it and how to overcome it; these questions are address in section 7.

5.3 Cost-effectiveness of energy efficiency upgrades

The cost-effectiveness of energy-conservation measures may be evaluated in different context, from a micro perspective (e.g. end-user) perspective to a macro perspective (e.g. energy supplier or government) perspective. In either case, there may be trade-offs and synergies that affect the cost-effectiveness of energy-conservation measures. In order to assess these, it is necessary to consider competing alternatives such as investments in the energy system.

In **Paper I**, the necessity of considering the cost-effectiveness of energy efficiency upgrades in a broader context was identified in several papers.

5.4 Sub-conclusion

In **Paper I**, a number of challenges that remain in building stock energy modelling were identified, highlighting the need for:

[15] The Rebound Effect describes the systematic response to increases in efficiency, e.g. changes in user behaviour

- flexible and transparent building stock energy models that are capable of emulating the variety in the building stock
- models that are capable of providing a building-physical description while accounting for user behaviour
- models that can be validated against actual consumption data
- models that can readily be integrated in other types of models in order to identify trade-offs (e.g. between investments in energy efficiency upgrades and investments in the energy supply system)

Improving the flexibility of BSEMs could facilitate identification of subsets that possess a cost-effective energy-saving potential.

Secondly, confidence in BSEM results rests upon the reliability of the results, which can only be ensured through model validation. This depends in turn on correct modelling of user behaviour, e.g. in order to account for rebound effects.

Lastly, investments are rarely isolated, for which reason it should be possible to consider investments in energy efficiency upgrades in combination with competing alternatives.

Thus, incorporating the aspects above could improve the reliability of BSEMs and make them more useful as political decision-making tools.

6 A sample-based bottom-up model of the Danish residential building stock

*Unless otherwise specified, this section is based on **Paper II***

With the introduction of the Energy Performance of Buildings Directive (EPBD), vast amounts of building-physical information has become available. So far as this is collected and stored disaggregated by component, the individual buildings' energy demand can be calculated anew, as suggested in **Paper A**.

In **Paper II**, data from the Danish EPC database was used for setting up a sample-based bottom-up model, in order to overcome the drawbacks of traditional archetype models discussed in section 5.

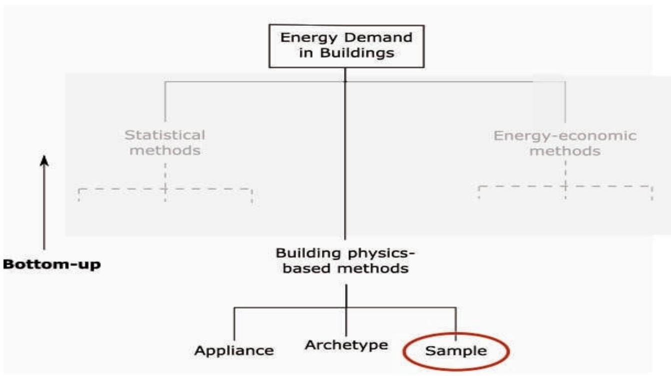


Fig. 10: A sample based bottom-up model

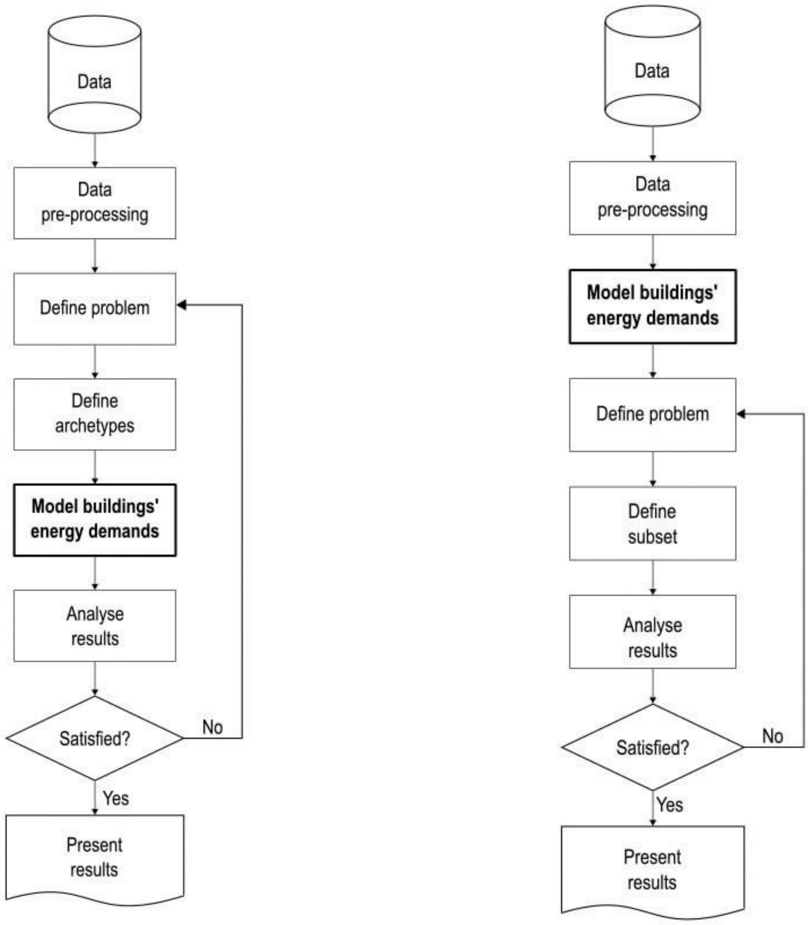
This approach provided a unique building-physical description (in terms of a monthly-mean heat balance) of each building in the database, offering a number of advantages including:

- Ideal model flexibility
- Preserved diversity of individual building's energy demand
- Possible validation of model results

6.1 Flexibility of an individual-building model

Identifying cost-optimal investments in energy efficiency upgrades entails considering various subsets of the building stock, e.g. buildings with different characteristics. Moreover, different segments of the building stock may be interesting to different stakeholders or in different (political) contexts. Therefore, a desirable feature of a BSEM is the ability to study different subsets, e.g. visualise the energy-saving potential, without having to fit a model specific to each subset. Therefore, segmenting the building stock before analysing energy demands and the related energy-saving potential is undesirable when the model is to be used as a tool for advising political decision-makers and other stakeholders.

Traditional building archetype based models are faced with the challenge of how to segment the building stock appropriately (i.e. in a way that minimises errors in the model). Moreover, these models remain inflexible, as separate models must be defined for each subset under investigation. On the other hand, sample based BSEMs, which considers each building in the stock individually, make it possible to study any segment of the building stock without defining a particular subset prior to modelling. Figure 11 illustrates the flexibility of a sample-based BSEM compared with an archetype-based BSEM conceptually.



(a) Archetype-based BSEM flowchart **(b)** Sample-based BSEM flowchart

Fig. 11: Conceptual illustration of the difference in workflow between an archetype-based BSEM and a sample-based BSEM

In traditional archetype-based BSEMs, a specific model must be fitted to each subset. Contrary, in a sample-based BSEM, all building's energy demands are modelled individually prior to studying a particular subset of the building stock. Therefore, modelling buildings' energy demands before (i.e. outside) the analysis- and evaluation loop ensures maximum flexibility of the sample-based BSEM.

6.2 Capturing building heterogeneity

Due to the diversity of the building stock, buildings' energy demands vary immensely. Even within groups of buildings with similar characteristics, there is a significant variation in energy demand. Defining archetypes (i.e. segmenting the building stock) in a way that minimises the variation in each segment remains a challenge. Calculating the energy demands of each building in a sample individually preserves the diversity in the output of the model contrary to traditional archetype-based BSEMs that provide point estimates of the energy demand in each segment of the building stock.

Figure 12 illustrates the heterogeneity in energy demand in detached single-family houses in relation to the year of construction. The horizontal lines illustrates the heat demand calculated in an archetype-based model, which segments the building stock into nine construction periods. It should be noted that these energy demands were calculated on the basis of the physical characteristics of the buildings alone, i.e. assuming the same indoor environmental conditions in all buildings. Hence, user influence and building specific environmental conditions could cause the observed variation to be even larger.

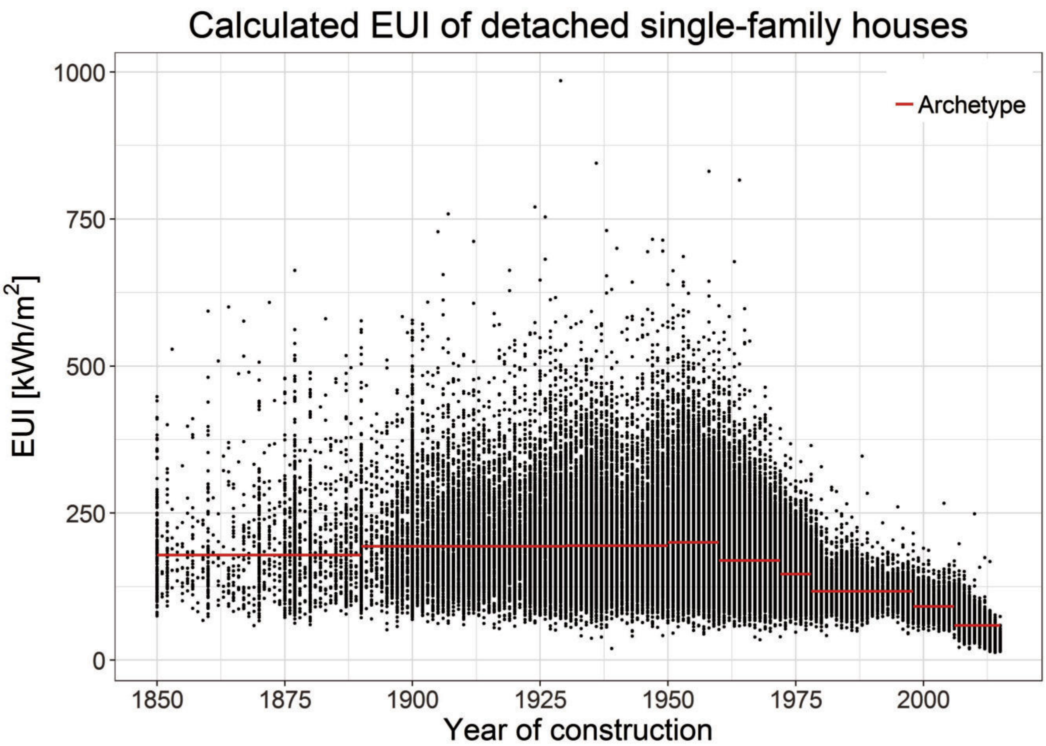


Fig. 12: Buildings' energy demand vary significantly (black dots) even within groups of buildings with similar characteristics. Using simple archetypes (red lines) to estimate buildings' energy demand entails significant loss of variation on the individual building level (adapted form **Paper B**).

In **Paper B**, the accuracy of using archetypes for estimating the average heat demand was assessed. In summary, the archetypes offered a nice emulation of the average heat demand (i.e. below 10 % in almost all cases). However, as the building archetypes only provided a coarse representation of the building stock, much variation is lost within each segment.

Identifying energy-inefficient buildings

From a building-physical point of view, the cost-effectiveness of employing an energy-conservation measure is closely linked with the energy efficiency of the building. Therefore, identifying energy-inefficient buildings represents a key step in building stock energy modelling.

As building archetypes only provide a point estimate of each segment of the building stock, it is not possible to identify energy inefficient buildings within each segment without access to consumption information (i.e. in a purely building-physics based setting). Therefore, very detailed archetypes, which minimise the variance in energy consumption, would be required in order to identify energy inefficient buildings.

On the other hand, energy-inefficient buildings can easily be identified in a sample-based BSEM, in which energy demands are calculated for each building individually. This makes it possible to target buildings with a particularly high energy demand to be targeted in an energy efficiency campaign. Likewise, subsequent analyses of the energy demand for particular energy end-uses can then be used for proposing appropriate energy-conservation measures.

The energy-efficiency (in terms of the energy-use intensity, EUI) of all buildings registered with an individual boiler was considered in **Paper II**. Figure 13 illustrates heat demand in this subset before and after implementation of a number of fictitious energy-conservation measures in a scenario analysis.

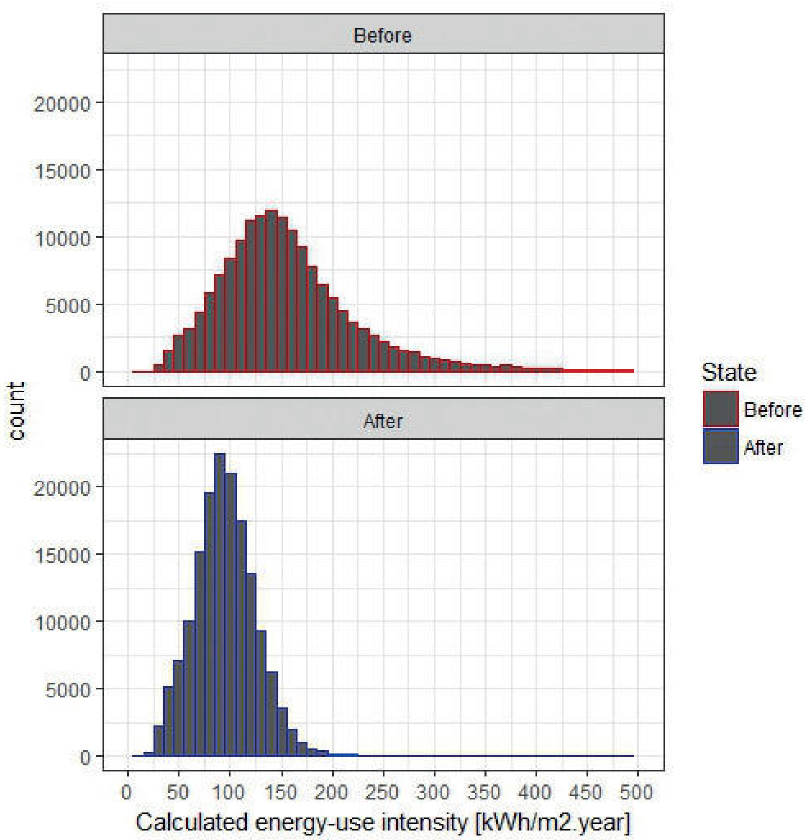


Fig. 13: Distribution of the heat demand in a particular subset of the building stock before and after a proposed energy-upgrading (scenario analysis) (**Paper II**)

Thus, a sample based BSEM provides the means for identifying energy-inefficient buildings as well as for studying effects of energy-upgrading these particular buildings.

6.3 Model validation

An important feature of the sample based model is the ability to link the output of the model with other information in order to assess the validity of the model.

In **Paper II**, this was exemplified by linking the calculated heat demands from the model with the corresponding heat demand from the BBR database. Table 4 illustrates the concept, in an example where a simple linear regression model was used for correcting the calculated energy-saving potential due to window replacement.

Model	Heat demand before [PJ]	Heat demand after [PJ]	Energy-savings [PJ]
Building-physical	161.7	138.0	23.7 (14.7 %)
Corrected	105.6	92.8	12.8 (12.1 %)

Table 4: Estimated energy-saving potential due to window replacement with and without correction (adapted from **Paper II**)

Due to the unique building representation, additional factors could be taken into account in order to show that estimates should be corrected differently depending on the characteristics of the building. This topic is treated in depth in section 7.

6.4 Sub-conclusion

Guiding policy-makers and other stakeholders with respect to the cost effectiveness of investments in energy-conservation measures requires a multi-detailed overview of the building stock. This includes an assessment of the energy-saving potential in various subsets of the building stock. Therefore, it is crucial that a building stock energy model provides the means for studying energy-efficiency in the building stock in a flexible and easy-to-communicate way.

By employing a sample-based BSEM the energy demand could be studied in an arbitrary subset of the building stock without pre-defining it in the model, thereby ensuring maximum flexibility.

Moreover, the variability of the energy demands in the building stock was preserved, allowing for identification of energy inefficient buildings to be targeted for energy-efficiency upgrades.

Finally, the model allowed for easy validation of model results, as individual buildings could readily be identified and linked with registered energy consumption.

One drawback of the sample-based modelling approach is the data requirement; considerable amounts of data must be available for this modelling approach to make sense. This data must be collected in a consistent manner, at a disaggregated level, in order to facilitate the array-programming approach used in **Paper II**. The data-intensity of the model could also cause it to become very computationally intensive if the level of detail in model was to be increased, e.g. if the temporal resolution was increased to include hourly time-steps. However, with a representative sample, smaller amounts of data (i.e. data on fewer buildings) could be sufficient to represent the building stock.

7 Hybrid bottom-up building stock energy modelling

*Unless otherwise specified, subsection 7.1 and subsection 7.2 are based on **Paper III** and subsection 7.3 is based on **Paper IV**.*

As suggested previously, assessing the energy-saving potential of a building stock requires a building-physical description of the buildings in question. However, user behaviour is often inappropriately described in traditional building-physics based models, thereby undermining the validity of the model.

This section discusses the use of hybrid BSEMs, which combine traditional building-physics based methods with statistical methods, as a method for improving estimates of the energy-saving potential in building stocks.

Building-physics based BSEMs provide the means to study the energy-saving potential due to an energy upgrade; however, a major drawback of building-physics based BSEMs is the need for defining user behaviour in terms of occupancy schedules, indoor temperature settings, air change rates, DHW consumption, etc. As this information is rarely available at the building stock level, normative values are often used in its place. However, this could compromise the validity of the model results. Moreover, several studies have suggested a rebound effect in relation to energy-upgrading of buildings; i.e. homeowners tend to improve the indoor environmental comfort when improving the energy performance of their house. Therefore, there is a need for models that makes it possible to account for these aspects.

By use of energy consumption data, statistical methods provide the means to study energy efficiency without explicitly modelling the users of the building. Therefore, in cases where energy consumption data is available, a natural extension of the traditional building stock energy modelling techniques would be to combine building-physics based methods and statistical methods into a hybrid BSEM. By leveraging on the advantages of each of the two methods, energy efficiency could be studied without explicitly modelling the users of the building. A hybrid BSEM could use statistical models as input in a building-physical model or vice versa. The hybrid BSEM concept is illustrated in Figure 14.

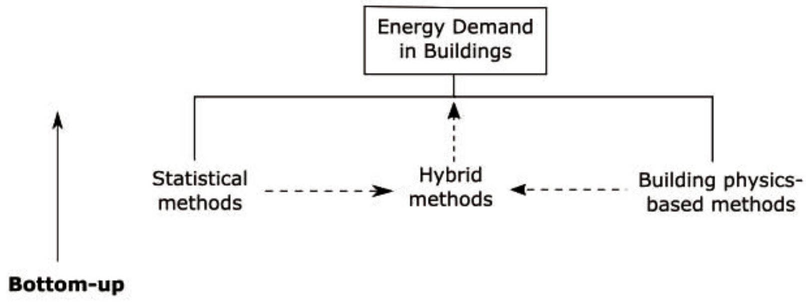


Fig. 14: Hybrid bottom-up building stock energy modelling (**Paper III**)

In **Paper III** and **Paper IV**, a hybrid BSEM that used the output of a building-physics based model (in term of the calculated heat demands) as input in a statistical (regression based) model was proposed. The workflow of this hybrid model type is illustrated in Figure 15.

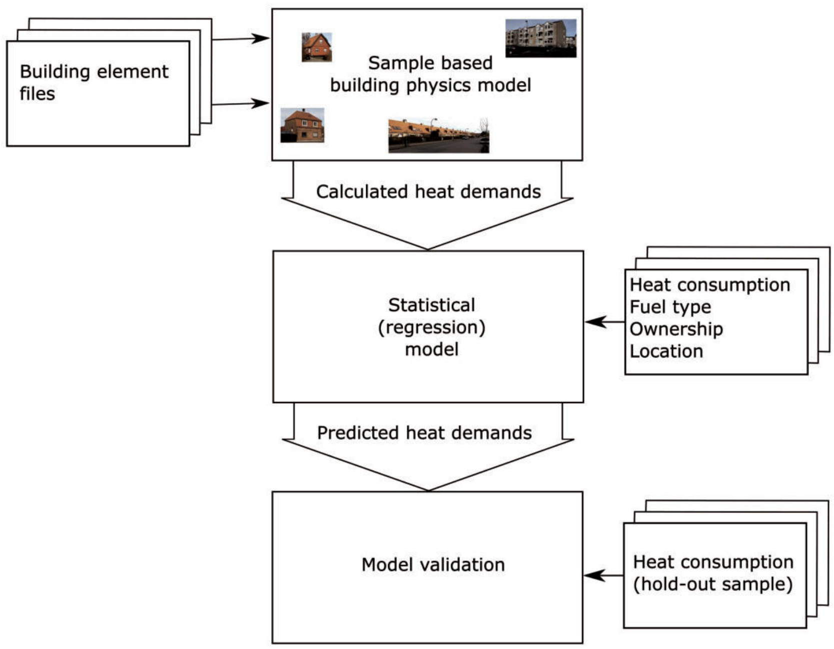


Fig. 15: Flowchart of a hybrid model that uses the output of a building-physics based model as input in a statistical model (**Paper III**)

It should be noted that the statistical model, including entering explanatory variables, was not identical in the two papers.

In the following sections we demonstrate how a hybrid BSEM of the Danish residential building stock could be used for obtaining more reliable results (i.e. improve the accuracy and validate the model), as well as for estimating the influence of rebound effects on the energy-saving potential in the building stock.

7.1 Technical- and realisable energy-saving potential of a building stock

Estimating the energy-saving potential in a building stock is a key feature of any BSEM. However, because the energy-saving potential does not only depend on the physics of the building, this is not a trivial task. Several studies have found that energy consumption that is estimated on the basis of the energy performance of the building alone (i.e. assuming normative indoor environmental conditions) is overestimated for buildings with a poor energy performance whereas it is underestimated for buildings with a good energy performance [20, 22, 24, 33].

In **Paper III** and **Paper IV**, a distinction was made between the *technical* and the *realisable* energy-saving potential (ESP). Whereas the technical ESP assumes fixed (normative) indoor environmental conditions, the realisable ESP takes rebound- and prebound effects into account. Figure 16 illustrates the difference between the technical- and the realisable energy-saving potential.

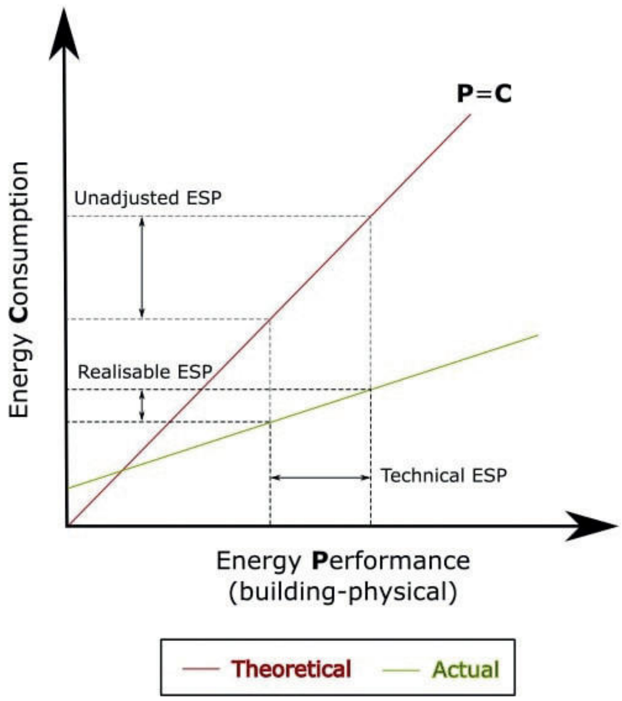


Fig. 16: Conceptual illustration of the difference between the technical- and the realisable energy-saving potential (**Paper III**)

This distinction was made in order to not overestimate the effect (i.e. the actual reduction in energy consumption) of an energy-efficiency upgrade on the energy consumption.

7.2 Improving the accuracy and validity of building stock energy models

Model validation was identified as a key to the credibility of BSEMs in in **Paper I**. **Paper III** was devoted to developing and optimising predictive performance of a hybrid BSEM, using available building stock data. Moreover, the proposed model was validated on the individual building level as well as on a building stock (i.e. aggregated) level.

Selecting a parsimonious model

With access to extensive amounts of information, it was necessary to determine which predictors (i.e. input variables) to include in the model. For this purpose, a model selection algorithm was developed with a view to optimising the predictive capabilities of the model while keeping it as simple as possible; i.e. selecting a parsimonious model. This algorithm is outlined in algorithm 1.

Data: Sub-sample using 50.000 observations ($\approx$ 40 % of all observations)
split in ten equally sized folds;
for *fold 1* **to** *10* **do**
    select fold *j* for cross-validation and fit base-model using the remaining nine folds;
    for *each additional predictor* **do**
        add predictor to the base model;
        predict energy consumption for the 10th fold (not used for fitting the model);
        calculate RMSE;
        select model with the lowest RMSE as new base model;
        return *order in which predictors were added and corresponding RMSEs*
    end
end

Algorithm 1: Model selection algorithm using Forward Subset Selection in combination with 10-fold cross-validation (**Paper III**)

Using a training sample to fit the model, each predictor (i.e. input variable) was added to the model to test the marginal improvement in prediction accuracy, i.e. Forward Stemwise Selection. In order to remedy any selection bias, ten-fold cross validation was employed. Finally, the parsimonious model was selected on the basis of a one standard error limit, see Figure 17.

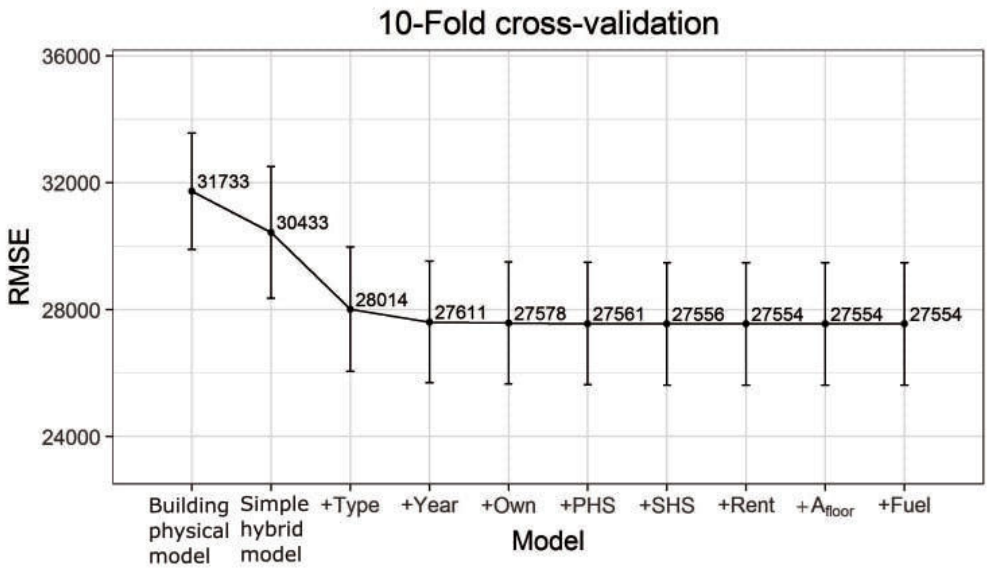


Fig. 17: Model selection based on Forward-Subset Selection and 10-fold cross-validation (**Paper III**)

Based on the available information, the best (parsimonious) prediction model was found to include the calculated heat demand (i.e. the building-physical model) and the building type as predictors in **Paper III**.

Model validation

In **Paper III**, four metrics were used for evaluating the fitted model; these were the explained variances (in terms of the adjusted R^2), the RMSE normalised by the registered heat consumption (CVRMSE), the mean absolute percentage error (MAPE) and the normalised mean bias error (NMBE). The choice of metrics, as well as their interpretation, was addressed in the paper.

The model was validated using the data that was not used for fitting the model (i.e. out-of-sample validation) in order to avoid overfitting. Table 5 summarises the model validation metrics used for evaluating the model fit in **Paper III**.

Model	R^2_{adj}	**CV(RMSE)**	**MAPE**	**NMBE**
Building-physics based model	75.5 %	121.6 %	51.1 %	-22.8 %
Hybrid model	81.1 %	106.8 %	31.9 %	-1.0 %

Table 5: Model validation based on four metrics (**Paper III**)

In all cases, the hybrid model outperformed the building-physics based model.

Especially in terms of the NMBE, the hybrid model improved the estimates of the heat consumption. This entails that almost all bias (i.e. on the aggregate level) was eliminated.

As the model included information about the building type, the predictive performance was also evaluated at this level in **Paper III**. This is shown in Table 6.

Model	**n**	$\mathbf{R}^2_{adj}$	**CV(RMSE)**	**MAPE**	**NMBE**
All buildings	40 217	81.1 %	106.8 %	31.9 %	-1.0 %
Farm houses	249	35.8 %	53.4 %	45.4 %	-2.5 %
Detached single-family house	29 304	37.0 %	35.7 %	30.1 %	1.2 %
Terraced houses	7258	54.6 %	57.4 %	37.2 %	4.1 %
Blocks of flats	3406	75.5 %	71.4 %	35.2 %	-4.5 %

Table 6: Hybrid model evaluation (model validation) considering all buildings collectively and each building type separately (**Paper III**)

7.3 Estimation of rebound effects

Rebound-effects have been proposed as the primary reason for the difference between the building-physical energy performance of a building and the corresponding energy consumption, i.e. the Performance gap [9, 23]. Studies suggest that occupants adjust their behaviour, in terms of indoor environmental comfort settings, according to the (building-physical) energy performance of the building. This is illustrated in Figure 18.

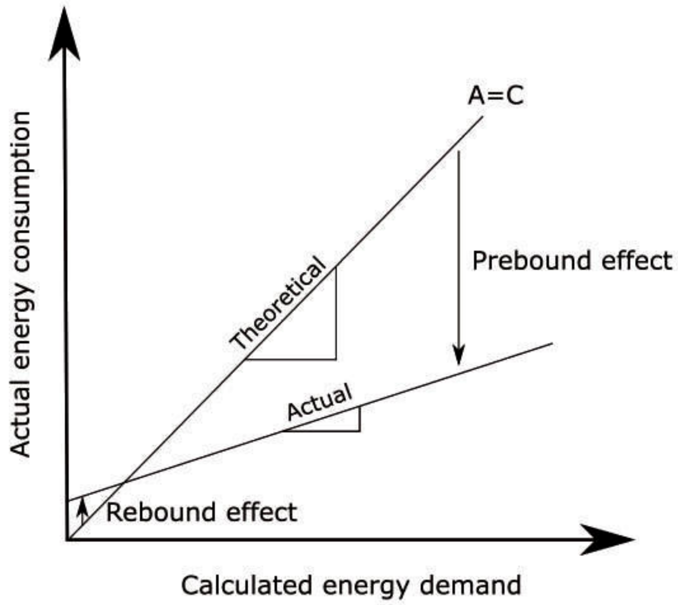


Fig. 18: Conceptual illustration of prebound- of rebound effects as an explanation for the performance gap (**Paper IV**)

Therefore, studying rebound effects in detail entails studying energy consumption in combination with indoor environmental conditions. However, on the building stock level, where access to indoor environmental conditions is often unavailable, other methods, which rely on available data, are needed for studying rebound effects.

By combining a building-physical description with energy consumption data in a hybrid BSEM, pseudo-rebound effects[16] could be studied, as suggested in **Paper IV**. Studying rebound effects in a hybrid BSEM has the additional advantages that available information could easily be incorporated. This way, differences in rebound effects between buildings with different characteristics were studied.

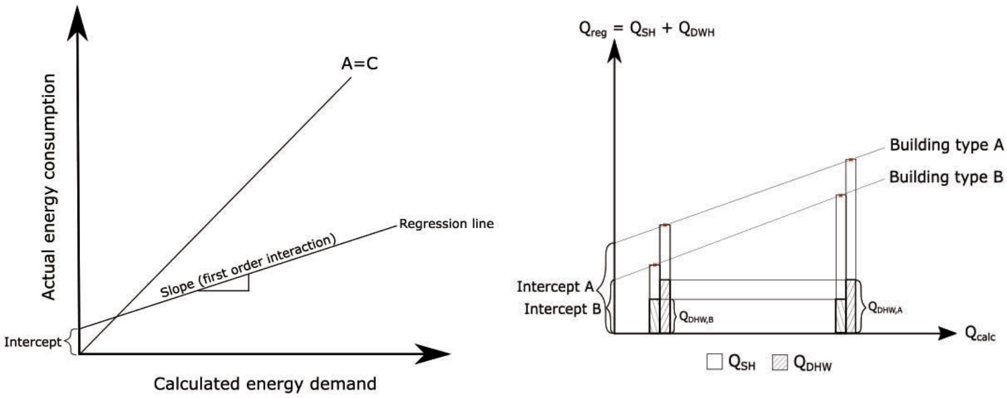
Model interpretation

In contrast to the prediction model described in subsection 7.2, focus was on the *interpretability* of the statistical model in **Paper IV**. Furthermore, emphasis was on the validity of the model in terms of not violating any model assumptions. Both interpretability and model validity are crucial in order to able to estimate the marginal effect of a given building characteristic (i.e. differences in rebound effects between buildings with different characteristics). Equation 1 outlines the statistical model used for estimating rebound effects in **Paper IV**.

[16]The term "pseudo" was used to indicate that energy consumption was not studied before and after an energy upgrade but rather as a function of the building-physical energy performance.

$$
\begin{aligned}
\log(Q_{\text{reg},i}) = {} & \beta_0 + \beta_1 \cdot \log(Q_{\text{calc},i}) \\
& + \beta_2 \cdot Type_i + \beta_3 \cdot PHP_i + \beta_4 \cdot SHS_i \\
& + \beta_6 \cdot Rent_i + \beta_7 \cdot Mun_i \\
& + \beta_8 \cdot Type_i \cdot \log(Q_{\text{calc},i}) + \beta_9 \cdot PHS_i \cdot \log(Q_{\text{calc}}, i) \\
& + \beta_{10} \cdot SHS_i \cdot \log(Q_{\text{calc},i}) + \beta_{12} \cdot Rent_i \cdot \log(Q_{\text{calc},i}) + \epsilon_i
\end{aligned} \tag{1}
$$

The intercepts of the model (i.e. the main effects, β_0 and β_2 through β_7) were used for accounting for the part of the heat consumption that was used for DHW preparation. Rebound effects were estimated in terms of the interaction effects (i.e. the parameter estimates β_1 and β_8 through β_{12}), as these can be interpreted as differences in slopes among the explanatory variables in the model, see Figure 19.



(a) Model interpretation: The slopes (i.e. the interaction terms) of the statistical model were used for estimating rebound effects

(b) In the statistical model, the intercepts (i.e. the main effects) were used for account the fraction DHW preparation constituted.

Fig. 19: Model interpretation (**Paper IV**)

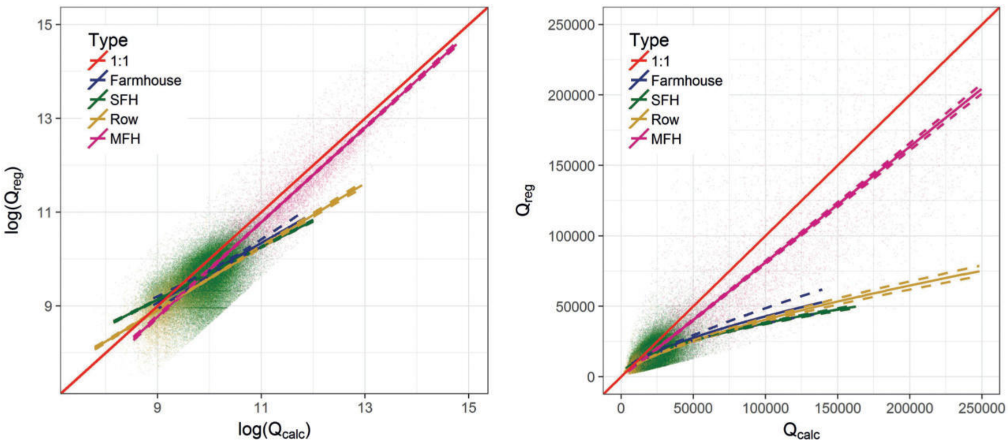
Table 7 lists the estimated slopes (including confidence intervals, CI) of the regression model.

Parameter	0.5 % CI	Estimate	99.5 % CI	Std. Error	p-value
$\log(Q_{calc})$	0.52	0.61	0.71	0.03	4.99×10^{-105}
$\log(Q_{calc})$:Type$_{SFH}$	−0.21	−0.12	−0.02	0.03	3.93×10^{-5}
$\log(Q_{calc})$:Type$_{Row}$	−0.11	−0.01	0.08	0.03	0.64
$\log(Q_{calc})$:Type$_{MFH}$	0.21	0.30	0.40	0.03	4.47×10^{-26}
$\log(Q_{calc})$:PHS$_{DH}$	0.04	0.05	0.07	<0.01	5.72×10^{-41}
$\log(Q_{calc})$:SHS$_{El}$	−0.03	0.01	0.05	0.01	0.28
$\log(Q_{calc})$:SHS$_{Stove}$	0.02	0.04	0.06	0.01	2.86×10^{-9}
$\log(Q_{calc})$:SHS$_{Both}$	−0.001	0.07	0.13	0.02	0.71×10^{-4}
$\log(Q_{calc})$:Rent$_{Yes}$	0.04	0.05	0.07	<0.01	2.92×10^{-40}

Table 7: Parameter estimates from the MLR model (**Paper IV**)

In most cases, a statistical significant difference was observed among buildings with different characteristics, indicating different degrees of rebound effects in buildings with different characteristics. Thus, the realisable energy-saving potential could be expected to depend on the building characteristics.

Taking the four different building types as an example, the estimated slopes could be illustrated as shown in Figure 20.



(a) Q_{calc} vs. Q_{reg} in the log-transformed variable space **(b)** Q_{calc} vs. Q_{reg} in the back-transformed variable space

Fig. 20: Relationship between Q_{calc} and Q_{reg} for the four building types. The dashed lines represent the respective confidence intervals (**Paper IV**)

Though potential causes for the observed differences were not analysed in depth, the analyses did warrant further investigation of the phenomena.

7.4 Sub-conclusion

A building-physical description is necessary in order to evaluate the energy-saving potential of a building stock due to an energy-efficiency upgrade. However, the use of normative assumptions, due to lack of data on the users of the buildings, renders purely building-physics based models inaccurate. In order to get an accurate estimate of the energy-saving potential in a building stock, user behaviour - as well as changes in user behaviour - must be modelled correctly.

A hybrid BSEM, which combines traditional a building-physics based model with a statistical model, provides the means for studying the energy-saving potential of a building stock without modelling the users of the building explicitly. Moreover, because a hybrid model includes energy consumption data, the validity of the model can be assessed in order to ensure trustworthy results.

In the present study, the output of a building physics-based model was used as input in a statistical model. This had several distinct advantages including:

- ✓ Energy-savings could be estimated by means of traditional engineering (thermodynamic) calculations
- ✓ User behaviour could be modelled implicitly by means of statistical methods
- ✓ Rebound effects could be estimated[17]
- ✓ The accuracy of the model could be evaluated

However, the proposed hybrid modelling approach was also subject to a few draw-backs; these were:

- − Specific aspects of user behaviour could not be identified

With the proposed model, it was possible to account for effects that were inherent to the (building-physical) energy performance of the building. This could provide more accurate estimates of the actual decrease in energy consumption following implementation of an energy-conservation measure, i.e. the realisable energy-saving potential.

[17] in order to estimate rebound effects, an interpretable model (i.e. a model in which the model estimates have physical meaning) must be used, e.g. a linear regression model

8 Estimation of the energy-saving potential of a building stock

*This section present an abstract of **Paper V**. As this paper is still only a draft, this section may be read independently of the paper. The content reflects only the author's contribution to the paper, as indicated by the co-author statement.*

As discussed in **Paper I**, investments in energy-upgrades should take the context into account in order to assess trade-offs between these investments and other investments, e.g. investments in the energy supply system. Therefore, the model should be integrable with other models in order to assess these trade-offs.

In **Paper V**, the hybrid building stock energy model (BSEM), which was introduced in section 7, was integrated in an economic model in order to evaluate the cost-effectiveness of energy-conservation measures from an end-user perspective, serving to illustrate the applicability of the model.

8.1 Identifying cost-optimal levels of energy-conservation measures

As the cost-effectiveness of an energy-conservation measure (ECM) is intimately linked with the energy performance of the building, it is crucial that the heterogeneity of the buildings' energy performance is captured by the BSEM. In addition, the costs of energy upgrading a building component is specific to the individual component, as sizes and component types vary. Therefore, both must be taken into consideration, when evaluating the cost-effectiveness of an energy upgrade.

In **Paper V**, this was done by means of cost-curves, which were developed for determining cost-optimal levels of five building components, namely roofs, external walls, floors, windows and mechanical ventilation. Figure 21 depicts the costs of reducing the heat demand in the considered building stock by a given amount by means of energy upgrading external walls (cavity walls and solid walls respectively).

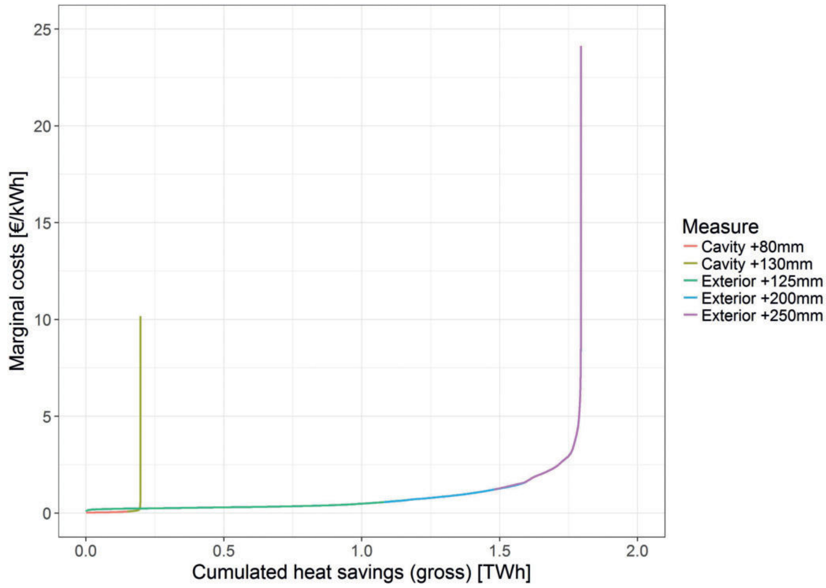


Fig. 21: Price of reducing the heat demand in a sample of 12.589 buildings by means of energy upgrading external walls in the building stock (**Paper V**)

The marginal costs of energy-upgrading each component was compared with the corresponding consumer price of district heating, in order to identify cost-optimal levels of each ECM in each building individually. In the paper this was used for studying effects of changing the district heating tariff structure.

8.2 Technical and realisable energy-saving potential

The cost-effectiveness of imposing the proposed ECMs was evaluated from a technical- as well as a realisable perspective. The technical energy-saving potential was assessed using the building-physics based model, whereas the realisable energy-saving potential was estimated using the hybrid BSEM developed in section 7.

Considering the marginal costs of realising the cost-effective energy-saving potential in the considered sample, it immediately appears that there was a significant difference between the technical (i.e. gross) energy-saving potential and the realisable (i.e. net) energy-saving potential, see Figure 22.

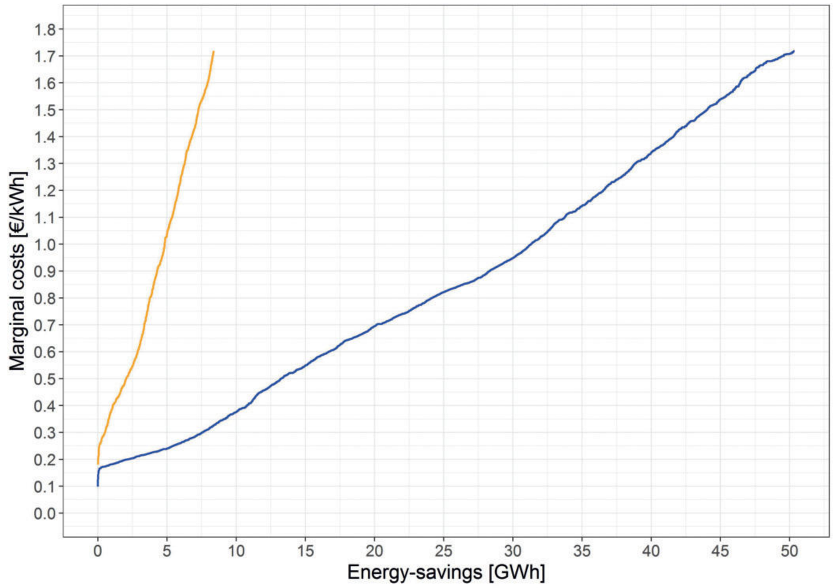


Fig. 22: Marginal costs of employing cost-effective ECMs in the considered subset of the Aarhusian building stock. The yellow line denotes the technical (gross) energy-savings whereas the blue denotes the realisable (net) savings (**Paper V**)

Table 8 lists the absolute and relative energy-saving potential, given an implementation of all cost-effective ECMs.

Table 8: Technical- and realisable energy-saving potential of all cost-effective ECMs in the considered subset of the Aarhusian building stock (**Paper V**)

Energy-saving potential	**Absolute** [MWh]	**Relative** [%]
Technical (gross)	50.4	9.3%
Realisable (net)	8.4	1.9%

Evidently, there was a significant difference between the technical- and the realisable energy-saving potential, as rebound effects could be expected to counteract the decrease in energy consumption, thereby causing less ECMs to be cost-effective. This could have major implications, not only for the building owner, but also bodies that subsidised the energy-upgrade (e.g. district heating companies or governmental subsidy programs) with the expectation that the energy consumption would decrease by an amount corresponding to the technical energy-saving potential.

8.3 Sub-conclusion

Estimating the energy-saving potential of a building stock often involves identification of cost-optimal levels of particular energy-conservation measures (ECMs). As the cost-effectiveness of an ECM is context dependent, it is necessary to take external conditions into account, e.g. by integrating the BSEM with other models.

In **Paper V**, the applicability of the BSEM was illustrated in a study of the cost-effectiveness of a number of energy-conservation measures from an end-user perspective, using a Danish district heating area as a case-study. Due to the disaggregated output of the proposed BSEM, it could easily be integrated with order models (e.g. an economic model as proposed in **Paper V**), in order to identify cost-optimal levels of different ECMs. Moreover, owing to the design of the BSEM, this could readily be extended to other/more areas, as well as be used in other contexts, e.g. in order to assess trade-offs between this investment in energy efficiency upgrades and investments in other alternatives (e.g. renewable energy technologies).

Conclusion

"The more I learn, the less I realise I know."
- Socrates

9 Discussion

In the present thesis, a sample-based hybrid bottom up building stock energy model of the existing Danish residential building stock was developed. This type of model offered a number of advantages, including superior flexibility and the ability to account for rebound effects. However, despite the progress made, many challenges remain. This section discusses some of the issues that were not accounted for in this study, which could be used for improving BSEMs in the future.

9.1 Data considerations

Limited access to data is a prerequisite in building stock energy modelling, for which reason existing data sources are often used. Therefore, assessments of the quality of existing data sets, which is questionable in many cases, constitute an important next step. Especially systematic errors should be assessed, e.g. systematic misjudgement of particular groups of buildings' thermal characteristics, as discussed in **Paper IV**. However, accessing the quality of data used is paramount in any case, especially if these have not been collected specifically for building stock energy modelling.

In addition, collecting information about the indoor environmental conditions across buildings with different characteristics could provide valuable information as to where current models are wrong. For instance, knowing the relationship between the average indoor temperature and the energy efficiency of buildings would be extremely valuable. Collecting high quality data, e.g. measured thermal characteristics and air change rates, for smaller samples of buildings could also provide valuable information. With this information, normative assumptions could be adjusted to reflect actual conditions in the building stock.

9.2 Uncertainties in building stock energy modelling

In the present study, uncertainties in the building-physical model were considered implicitly on a building level, as described in section 7. Therefore, a natural extension of the model would be to consider uncertainties at a less aggregated level. This could include sensitivity analyses of the input parameters, as well as extending the model to include probabilistic (i.e. stochastic) inputs, as suggested by Booth et al. [3].

Building stock heterogeneity

In addition to the heterogeneity of the building-physical characteristics, addressed in subsection 6.2, heterogeneity also occurs in environment- (e.g. different buildings are exposed to different weather) and people (i.e. the users of buildings are different from one building to another) related data. Therefore, a natural extension of the model would be to use building-specific weather and people data. However, so far as this is available, the design of the sample-based BSEM allows for doing so readily.

Temporal resolution

Using a temporally aggregated model, in this case a monthly mean, could disguise errors that even out on average, e.g. differences between occupied- and non-occupied hours. Moreover, there are likely differences in consumption profiles, between buildings with different characteristics, e.g. single-family houses and multi-family houses. Increasing the temporal resolution of the model could help resolve some of these issues, as also noted by Gionniou et al. [21]. Increasing the temporal resolution of the model would also enable additional analyses, e.g. of peak demands in an energy system.

Model validation

Model validation is a key element in ensuring the reliability of a BSEM. In addition to the methods used for model validation in this thesis, cross-model comparisons (e.g. across countries) could provide valuable information about uncertainties that were not accounted for in a particular BSEM, e.g. aspects that were not included in a model (i.e. ontological uncertainties). However, cross-model comparisons are often complicated by different applications of models, as well as use of different data.

Finally, validating BSEMs that analyse the energy consumption over time, i.e. scenario analyses, is challenging because of changing household compositions (e.g. people moving in and out of houses). Failing to address this issue could deteriorate the accuracy of BSEMs in terms of increased uncertainty over time.

9.3 Extrapolation

Lastly, the models in this paper considered a sample of the building stock, for which data was available. Therefore, analysing the entire building stock would require extrapolation of the results, in order to cover the all buildings, i.e. spatial extrapolation. This could be done using data that is available on the building stock level, e.g. the heated floor area which is often available in public registries. In **Paper II**, this was done by means of simple linear regression. This approach could be refined by including additional available information in the regression model.

In addition, the building stock was considered from a static point of view (i.e. a point-in-time perspective). This entails that forecasting (e.g. in terms of scenario analyses) would also require temporal extrapolation. In addition to considering the energy efficiency of the existing building stock, extrapolating results in time would require forecasting of rates of demolition and new built (i.e. stock dynamics), as well as other trends, e.g. increases in heated floor area per person. Likewise, it would be necessary to forecast energy prices and weather conditions in order to assess the cost-effectiveness of improving energy efficiency under changing climate conditions (e.g. increasing temperatures).

As discussed in **Paper I**, the considered sample of buildings must be representative of the building stock, in terms of the energy consumption, in order to get valid results. In the present thesis, data from the Danish EPC database was used. This entails that only building that had been certified were considered. However, as the certification of buildings is not random, the considered sample could be biased (i.e. not representative of the entire Danish building stock).

A representative sample of the building stock could be obtained by means of random sampling or stratified random sampling, sampling from all buildings in the stock. Using stratified random sampling, the characteristics that turned out to affect energy consumption most in **Paper IV** could be used for defining stratification parameters. Sampling from the entire building stock would require that data be collected for buildings for which data is not available (e.g. buildings that have not been certified). On the other hand, with a representative subset of the building stock available, much less data may be required in the modelling process.

10 Conclusion

A building stock often comprises tens of thousands of buildings with different building-physical characteristics. Moreover, the use, as well as the users, vary tremendously among buildings. Finally, differences in environmental conditions among buildings, e.g. in terms of wind exposure or local heat island effects, makes a building stock a complex character. Due to this complexity and the limited access to relevant data, getting an overview of the energy efficiency, as well as the related energy-saving potential of a building stock is not trivial. However, building stock energy models (BSEMs) provide the means to encapsulate the complexity of the building stock, if set up appropriately. This way, pathways for improving the energy efficiency of the building stock in a cost-optimal way can be communicated to policy makers and other stakeholders.

Multiple types of BSEMs exist, each with advantages and disadvantages. Building-physics based models provide the means to study the energy efficiency of the building envelope, as well as the technical systems/building services. However, getting an accurate estimate of buildings' energy consumption is complicated by individual user behaviour. Therefore, models that can account for both are required for getting an accurate estimate of the energy-saving potential of the building stock. Moreover, flexible models are required for encapsulating the diversity of the building stock, in order to identify subsets of the building stock that offer a cost-effective energy-saving potential. Lastly, model validation is paramount to ensure the reliability of the model.

Using disaggregated building-physical data from the Danish Energy Performance Certificate (EPC) database, a sample-based bottom-up building stock energy model was set up for modelling the energy demand of each building in the sample. This method offered superior emulation the complexity of the building stock. Moreover, this type of model was exceeding versatile in terms of studying different subsets of a building stock, which could be used for informing different stakeholders.

This individual building-physical description of each building in the stock was combined with energy consumption data in a hybrid bottom-up model, which made it possible to account for user-behaviour implicitly in order to account for rebound effects. This way, the accuracy of the model could be improved substantially in order to provide reliable estimates of the energy-saving potential of the building stock.

Finally, the proposed building stock energy model allowed for ready integration with other models in order to assess trade-offs between different investments in different contexts.

On this background, it can be concluded that with access to disaggregated building-physical- and energy consumption data, a versatile and accurate building stock energy model can be set-up for assessing the energy-saving potential of a building stock. This type of model makes it possible to encapsulate the complexity (e.g. in terms of the heterogeneity) of a building stock as well as to study subsets of the building stock without fitting models specifically to each subset. With this information, policymakers and other stakeholders can make informed decisions that will reduce energy consumption in buildings, thereby mitigating climate changes.

References

[1] ASHRAE, *2013 ASHRAE Handbook: Fundamentals*. ASHRAE Research, 2013. [Online]. Available: https://app.knovel.com/web/toc.v/cid:kpASHRAEC1/viewerType:toc/root_slug:2013-ashrae-handbook

[2] Boligministeriet, *Bygningsreglement 1977*. Boligministeriet, 1977. [Online]. Available: http://w2l.dk/file/502102/br_syvoghalvfjerds.pdf

[3] A. T. Booth, R. Choudhary, and D. J. Spiegelhalter, "Handling uncertainty in housing stock models," *Building and Environment*, vol. 48, no. 1, pp. 35–47, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.buildenv.2011.08.016

[4] E. Concerted Action, *2016 – Implementing the Energy Performance of Buildings Directive (EPBD)*, E. Maldonado, Ed. Lisbon: Concerted Action, 2016. [Online]. Available: http://www.buildup.eu/en/practices/publications/2016-implementing-epbd-featuring-country-reports-0

[5] Danish Energy Agency (Energistyrelsen), "Bekendtgørelse om ændring af bekendtgørelse om Håndbog for Energikonsulenter (HB2016) - retsinformation.dk," 2016. [Online]. Available: https://www.retsinformation.dk/Forms/R0710.aspx?id=201570

[6] ——, "Energistatistik 2016 (In Danish)," Danish Energy Agency, Tech. Rep., 2016.

[7] ——, "Kvalitetskontrol for energimærkninger fra 2017," Danish Energy Agency, Tech. Rep., 2017.

[8] E. G. Dascalaki, K. Droutsa, A. G. Gaglia, S. Kontoyiannidis, and C. a. Balaras, "Data collection and analysis of the building stock and its energy performance—An example for Hellenic buildings," *Energy and Buildings*, vol. 42, no. 8, pp. 1231–1237, 2010. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0378778810000459

[9] P. de Wilde, "The gap between predicted and measured energy performance of buildings: A framework for investigation," *Automation in Construction*, vol. 41, pp. 40–49, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092658051400034X

[10] U. D. o. E. (DOE), "Weather Data | EnergyPlus." [Online]. Available: https://energyplus.net/weather

[11] DTU Management, "SAVE-E Energy Savings: Closing the Energy Efficiency Gap." [Online]. Available: https://www.sustainability.man.dtu.dk/english/research/eer/research-projects/save-e

[12] EPISCOPE, "Monitor Progress Towards Climate Targets in European Housing Stocks Main Results of the EPISCOPE Project - Final Project Report - (Deliverable D1.2)," Institute Wohnen und Umwelt GmbH, Tech. Rep. March 2016, 2016. [Online]. Available: http://episcope.eu/monitoring/overview/

[13] European Commision, "Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings (recast)," pp. 13–35, 2010.

[14] E. European Commision, “Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012 on energy efficiency,” *Official Journal of the European Union Directive*, no. October, pp. 1–56, 2012.

[15] European Commission, “2020 climate & and energy package | Climate Action.” [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2020_en

[16] ——, “2030 climate & and energy framework | Climate Action.” [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2030_en

[17] ——, “2050 low-carbon economy | Climate Action.” [Online]. Available: https://ec.europa.eu/clima/policies/strategies/2050_en

[18] ——, “Paris Agreement | Climate Action.” [Online]. Available: https://ec.europa.eu/clima/policies/international/negotiations/paris_en#tab-0-0

[19] Europen Commission, “Buildings - European Commission.” [Online]. Available: https://ec.europa.eu/energy/en/topics/energy-efficiency/buildings

[20] R. Galvin, *The Rebound Effect in Home Heating*. Routledge, Taylor & Francis Group, 2016.

[21] P. Gianniou, A. Heller, P. S. Nielsen, and K. Negendahl, “AGGREGATION OF BUILDING ENERGY DEMANDS FOR CITY-SCALE MODELS Department of Civil Engineering , DTU , Lyngby , Denmark Department of Management Engineering , DTU , Roskilde , Denmark,” *CISBAT 2015 International Conference*, 2015. [Online]. Available: http://cisbat.org

[22] K. Gram-hanssen, “How building regulations ignore the use of buildings , what that means for energy consumption and what to do about it,” in *ECEEE Summer Study Proceedings*, Hyéres, France, 2017, pp. 2095–2104.

[23] K. Gram-Hanssen and S. Georg, “Energy performance gaps: promises, people, practices,” *Building Research and Information*, vol. 46, no. 1, pp. 1–9, 2018. [Online]. Available: https://doi.org/10.1080/09613218.2017.1356127

[24] R. Haas, H. Auer, and P. Biermayr, “The impact of consumer behavior on residential energy demand for space heating,” *Energy and Buildings*, vol. 27, no. 2, pp. 195–205, 1998. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0378778897000340

[25] IEA EBC, “Building Energy Performance Assessment Based on In-situ Measurements (IEA Annex 71).” [Online]. Available: http://iea-ebc.org/projects/project?AnnexID=71

[26] IEA-EBC, “IEA Annex 70 Data Registry,” 2018. [Online]. Available: https://iercdevsrv.tyndall.ie/iea/testing/

[27] Intergovernmental Panel on Climate Change (IPCC), “Climate Change 2014 Synthesis Report Summary Chapter for Policymakers,” International Panel on Climate Change (IPCC), Tech. Rep., 2014. [Online]. Available: https://www.ipcc.ch/report/ar5/syr/

[28] ——, “Global warming of 1.5 °C,” Intergovernmental Panel on Climate Change (IPCC), Tech. Rep. June, 2018. [Online]. Available: https://www.ipcc.ch/sr15/

[29] M. Kavgic, A. Mavrogianni, D. Mumovic, A. Summerfield, Z. Stevanovic, and M. Djurovic-Petrovic, "A review of bottom-up building stock models for energy consumption in the residential sector," *Building and Environment*, vol. 45, no. 7, pp. 1683–1697, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360132310000338

[30] C. Kern-hansen, "Technical Report 13-19 2001 – 2010 Danish Design Reference Year - Reference Climate Dataset for Technical Dimensioning in Building , Construction and other Sectors Peter Grunnet Wang Kim Bjarne Wittchen Colophon," Danish Ministry of Climate, Energy and Building (DMI), Tech. Rep., 2013.

[31] J. Kragh and K. Wittchen, "Development of two Danish building typologies for residential buildings," *Energy and Buildings*, vol. 68, pp. 79–86, 2014. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S037877881300604X

[32] B. Lin, "A PROTOCOL-BASED, INVERSE-MODEL-DRIVEN METHODOLOGY FOR BUILDING AUDITING AND FORWARD ENERGY MODEL FORMULATION," Ph.D. dissertation, The Pennsylvania State University, The Graduate School, 2015.

[33] D. Majcen, L. Itard, and H. Visscher, "Actual and theoretical gas consumption in Dutch dwellings: What causes the differences?" *Energy Policy*, vol. 61, pp. 460–471, 2013. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S030142151300503X

[34] Ministry of Housing Communities & Local Government, "English Housing Survey: Energy efficiency 2016," 2018. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/724339/Energy_efficiency_2016.pdf

[35] C. S. Monteiro, C. Costa, A. Pina, M. Y. Santos, and P. Ferrão, "An urban building database (UBD) supporting a smart city information system," *Energy and Buildings*, vol. 158, pp. 244–260, 2018. [Online]. Available: http://dx.doi.org/10.1016/j.enbuild.2017.10.009

[36] NASA, "Global Climate Change - Causes." [Online]. Available: https://climate.nasa.gov/causes/

[37] N. Oreskes, "BEYOND THE IVORY TOWER: The Scientific Consensus on Climate Change," *Science*, vol. 306, no. 5702, pp. 1686–1686, 12 2004. [Online]. Available: http://www.sciencemag.org/cgi/doi/10.1126/science.1103618

[38] C. f. E. E. RCUK, "IEA-EBC Annex 70 – Building Energy Epidemiology: Analysis of Real Building Energy Use at Scale." [Online]. Available: https://energyepidemiology.org/

[39] C. F. Reinhart and C. C. Davila, "Urban building energy modeling e A review of a nascent field," *Building and Environment*, vol. 97, pp. 196–202, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.buildenv.2015.12.001

[40] G. Sousa, B. M. Jones, P. A. Mirzaei, and D. Robinson, "A review and critique of UK housing stock energy models, modelling approaches and data sources," *Energy and Buildings*, vol. 151, pp. 66–80, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.enbuild.2017.06.043

[41] B. Street and N. Scotia, "A NEW HYBRID END-USE ENERGY AND EMISSIONS MODEL OF THE CANADIAN HOUSING STOCK Department of Mechanical and Aerospace Engineering , Carleton University , 1125 Colonel By Drive , Ottawa , Ontario , K1S 5B6," pp. 109–116, 2008.

[42] L. Swan, V. I. Ugursal, and I. Beausoleil-Morrison, "IMPLEMENTATION OF A CANADIAN RESIDENTIAL ENERGY END-USE MODEL FOR ASSESSING NEW TECHNOLOGY IMPACTS Department of Mechanical Engineering , Dalhousie University , Department of Mechanical and Aerospace Engineering , Carleton University , 1125 Colonel By Dri," *Eleventh International IBPSA Conference*, vol. 11, no. Building Simulation 2009, pp. 1429–1436, 2009.

[43] L. G. Swan and V. I. Ugursal, "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 8, pp. 1819–1835, 2009.

[44] TABULA Project Team, "Typology Approach for Building Stock Energy Assessment - Main Results of the TABULA project," Institute Wohnen und Umwelt GmbH, Tech. Rep. June 2009, 2012. [Online]. Available: https://ec.europa.eu/energy/intelligent/projects/sites/iee-projects/files/projects/documents/tabula_final_report_en.pdf

[45] B.-o. B. Trafik, "BR18 - Energiforbrug." [Online]. Available: http://bygningsreglementet.dk/Tekniske-bestemmelser/11/Krav/250_256

[46] UNHCR, "UNHCR - Climate Change and Disasters." [Online]. Available: http://www.unhcr.org/climate-change-and-disasters.html

Appendix 1: Data exclusion

The following appendix contains a list of criteria used for excluding data that were deemed faulty. This appendix was provided in order to give the reader an opportunity to assess the reasonableness of the imposed criteria, as well as to justify the the number of buildings that were excluded from the analyses. Whereas most characteristics should be self-explanatory, others may be inherent to Danish calculation procedures.

It should be noted that the same building could be deemed faulty in several regards (e.g. having no windows and no heating system), in which case it would count in the *'No. of buildings'* column in more than one table. Therefore, the total number of buildings that were excluded did not necessarily equal the sum of all totals across tables. Moreover, this appendix only lists criteria that applied to the information used in the building-physical model, e.g. information building about the building envelope and heat distribution system, etc. In the actual data exclusion process, additional filters were imposed, e.g. criteria related to supply system (PV, heat pumps, etc.) in order to prepare this data for future inclusion in the model.

The structure of the appendix follows the structure of the Danish EPC database, which again follows the structure of the Danish EPC tools. Therefore, the tables in each section should give an indication of the information that was available.

Building information

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Heat capacity [$\frac{\mathrm{Wh}}{\mathrm{m}^2\cdot\mathrm{K}}$]	$[40;160]$	4219	4219
Hours of occupation [h/week]	[0;168]	353	353
Building type [-]	valid	301	301
Addition to energy frame [$\frac{\mathrm{kWh}}{\mathrm{m}^2\cdot\mathrm{year}}$]	≥ 0	10	10
Total		4806	4806

Table 9: Criteria for- and number of faulty registrations in the table containing building envelope elements in the Danish EPC database.

Building envelope information

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Area [m^2]	> 0	18 786	6971
U-value [$\frac{\mathrm{W}}{\mathrm{m}^2\cdot\mathrm{K}}$]	$]0.03;7]$	1545	1286
Temperature factor [-]	$[0;1]$	12 160	11 273
Total		32 124	19 125

Table 10: Criteria for- and number of faulty registrations in the table containing building envelope elements in the Danish EPC database.

Linear thermal transmittance information

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Length [m]	> 0	5343	4300
Heat loss [$\frac{W}{m \cdot K}$]	> 0.1	117	116
Temperature factor[18] [-]	$[0;1]$ $[0;3]$ $[0;4.5]$	12 160	11 273
Total		32 124	19 125

Table 11: Criteria for- and number of faulty registrations in the table containing information about linear thermal transmittance in the Danish EPC database.

Window information

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Number [-]	>0	15 691	6818
Area [m^2]	> 0	10 917	5443
Inclination [°]	[0;90]	568	346
U-value [$\frac{W}{m^2 \cdot K}$]	$]0.2;7]$	3963	1211
Temperature factor [-]	$[0;1]$	419	236
Fraction of glazing [-]	[0;1]	146	106
Solar transmittance (g-value) [-]	[0;1]	74	47
Solar shading factor [-]	[-1;1]	371	228
Shading angle horizon [°]	[0;90]	1509	736
Shading angle eaves [°]	[0;90]	1661	892
Shading angle left [°]	[0;90]	1663	938
Shading angle right [°]	[0;90]	1564	910
Shading angle window hole [°]	[0;90]	386	272
Total		24 228	11 126

Table 12: Criteria for- and number of faulty registrations in the table containing window information in the Danish EPC database.

Ventilation information

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Area [m^2]	> 0	3953	2595
Time of operation [-]	[0;1]	4897	2339
Natural ventilation (winter) [$\frac{1}{s \cdot m^2}$]	≥ 0	105	97
Total ventilation (winter) [$\frac{1}{s \cdot m^2}$]	≥ 0.3	7864	6465
Heat recovery [-]	[0;1]	565	397
Inlet temperature [°C]	-18, 0, 18	2587	1591
SEL [$\frac{kJ}{m^3}$]	[0.2;6]	4126	3759
Total		25 162	16 411

Table 13: Criteria for- and number of faulty registrations in the table containing ventilation information in the Danish EPC database.

Internal heat gains

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Area [m^2]	> 0	1654	1596
Heat load from persons [$\frac{W}{m^2}$]	]0;10]	714	629
Heat load from appliances [$\frac{W}{m^2}$]	]0;16]	494	401
Total		2570	2402

Table 14: Criteria for- and number of faulty registrations in the table containing information about internal heat loads in the Danish EPC database.

Heat distribution system

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Supply temperature [°C]	[30;90]	76 629	76 379
Return temperature [°C]	[15;90]	76 659	76 409
T_{supply}-T_{return} [°C]	≥ 0	227	227
Total		77 096	76 846

Table 15: Criteria for- and number of faulty registrations in the table containing information about the heating distribution system in the Danish EPC database.

In this table, extremely many systems were registered with a supply- and/or return temperature of zero.

Heat distribution pipes

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Length [m]	> 0	9468	9087
Heat loss [$\frac{W}{m \cdot K}$]	[0;99]	4718	4553
Temperature factor [-]	$[0; 1]$	2458	2395
Total		10 047	9632

Table 16: Criteria for- and number of faulty registrations in the table containing information about heating distribution pipes in the Danish EPC database.

Domestic hot water

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Average consumption $[\frac{l}{m^2 \cdot year}]$	[0;300]	21 761	21 735
Temperature [°C]	≥ 20	546	543
Total		21 931	21 905

Table 17: Criteria for- and number of faulty registrations in the table containing information about domestic hot water in the Danish EPC database.

Domestic how water tanks

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Number [-]	≥ 0	71 703	71 675
Volume [l]	≥ 0	3024	3024
Share of DHW [-]	[0;1]	3025	3026
Supply temperature [°C]	[30;90]	22 432	22 400
Heat loss $[\frac{W}{K}]$	>0	49 405	49 106
Temperature factor [-]	[0;1]	3097	3097
Total		96 436	96 101

Table 18: Criteria for- and number of faulty registrations in the table containing information about domestic hot water tanks in the Danish EPC database.

Domestic hot water pipes

Characteristic	Criterion	No. of faulty registrations	No. of buildings excluded
Length [m]	> 0	9468	9087
Heat loss $[\frac{W}{m \cdot K}]$	[0;99]	4718	4553
Temperature factor [-]	$[0;1]$	2458	2395
Total		10 047	9632

Table 19: Criteria for- and number of faulty registrations in the table containing information about domestic hot water pipes in the Danish EPC database.

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-380-8

AALBORG UNIVERSITY PRESS